Assessing the Accuracy of Telluric Corrections

Allyson C. Sheneman\textsuperscript{1,2,3,4}

Advisors: Adam B. Langeveld\textsuperscript{1}, Amy Bartholomew\textsuperscript{2}

\textsuperscript{1}Center for Astrophysics and Planetary Science, Cornell University
\textsuperscript{2}Department of Physics and Astronomy, State University of New York at New Paltz
\textsuperscript{3}Honors Program, State University of New York at New Paltz
\textsuperscript{4}Department of Biology, State University of New York at New Paltz

Honors Thesis & Physics Senior Project
May 17, 2024
Abstract

Observing transiting exoplanets with ground-based telescopes and high-resolution spectrographs enables the resolution of individual absorption lines in the exoplanet transmission spectra. However, observing from the ground inherently introduces telluric contamination: spectral contamination from absorption due to molecules in the Earth’s atmosphere. We take high-resolution observations of a transiting exoplanet around a bright A-type star as a case study and use synthetic telluric molecfit models to remove contamination from water and oxygen molecules in Earth’s atmosphere. The quality of the telluric corrections was statistically assessed for several different telluric regions based on absorption depth and molecular absorption species. We find that corrections for shallow telluric lines are more robust than deeper telluric lines, though both depend similarly on airmass. Corrections for different molecular bands varied by region. Some regions demonstrate a higher dependency on airmass, potentially due to the wavelength, depth, or quantity of telluric lines. Finally, we determine that the most accurate corrections are performed at observations with airmass under 1.07 corresponding to a zenith angle of approximately 20.84 degrees. Whilst this is a somewhat limited airmass range, these results highlight the need for improving telluric models for future searches of water and oxygen features in Earth-like exoplanet transmission spectra with 30m-class telescopes. These results may assist in optimizing observations for retrieving and preserving more data in an exoplanet transmission spectrum, especially when absorption features from the same molecules—and potential biosignatures—in Earth-like atmospheres fall in these highly contaminated regions.

Key Words: astronomy; astrobiology; exoplanets; atmosphere; spectroscopy; telluric contamination; telluric corrections
Acknowledgements

I am deeply grateful to my esteemed advisors, Dr. Adam Langeveld and Dr. Amy Bartholomew, for their continuous support, guidance, and discussions throughout the research process. Their expertise and encouragement have played a pivotal role in shaping the trajectory of this work. I also extend my heartfelt appreciation to the dedicated individuals of Dr. Ray Jayawardhana’s research team at Cornell University, notably Dr. Laura Flagg and Dr. Jake Turner, who welcomed me into their group and provided contributions and insights that have immensely enhanced this project.

I am profoundly grateful to Cornell University for graciously hosting this project through the Astrophysics and Planetary Science Research Experience for Undergraduates (REU) program during the summer of 2023. The support and resources provided by Cornell University, along with funding from the National Science Foundation (NSF), have been invaluable in facilitating this research endeavor and enriching my undergraduate research portfolio.

Furthermore, I thank the State University of New York at New Paltz, particularly the Physics & Astronomy Department and the Honors Program, for their steadfast support in sustaining this research during the academic year following the summer program as my Physics Senior Project and Honors Thesis, allowing for the culmination of this capstone endeavor.

Additionally, I am grateful to the American Astronomical Society for awarding me the Chambliss Astronomy Achievement Award at the 243rd annual conference for my poster design and presentation on this research. I would also like to thank the committee of the Astrobiology Graduate Conference (AbGradCon) hosted by Cornell University for accepting my attendance to present this research in June 2024.

Lastly, I would be remiss not to mention my peers, friends, family, colleagues, employers, and all other contributors to my success for their unwavering encouragement and belief in me. Thank you all for your immeasurable support in this journey.

REU Site Information: Astrophysics and Planetary Science Research at Cornell University. Primary Investigators Alexander Hayes and Britney Schmidt. Grant number: AST-2244064
1 Introduction

Since the groundbreaking first detection of an exoplanet atmosphere over two decades ago (Charbonneau et al., 2002 [3]), the field has experienced exponential growth. Six years following this milestone, the first ground-based observation of an exoplanet atmosphere was made, capable of resolving the sodium doublet (Redfield et al., 2008 [10]). As of 2021, nearly fifty distinct chemical species have been identified in exoplanet atmospheres and just under one hundred atmospheres have been examined (Lira-Barria et al., 2021 [9]).

High-resolution ground-based spectroscopy is utilized across disciplines and is an essential tool for numerous fields including exoplanet atmospheric spectroscopy (D. Shulyak et al., 2019 [13]), climate-induced atmospheric changes (H. Shi et al., 2023 [12]), and galactic observations (A. Alonso-Herrero et al., 2020 [1]; C. Conroy et al., 2019 [5]). Most exoplanet observations are made with ground-based instruments, which present many challenges. The primary challenge is telluric contamination: spectral contamination from the Earth’s atmosphere. The correction and removal of telluric lines proves difficult due to several factors, including their interaction with stellar lines and overlapping wavelength signals and noise (Snellen et al., 2008 [16]).

Telluric contamination, inherent to high-resolution ground-based spectroscopy, distorts spectral signals from celestial bodies. This interference is not uniform across the spectra and is dependent on atmospheric parameters, such as airmass, pressure, humidity, and temperature, which can alter the intensity of the telluric lines. Due to the dynamic nature of these atmospheric features throughout observation periods, it is difficult to accurately and consistently remove telluric lines. This is particularly pertinent for achieving individual line resolution in exoplanetary atmospheric spectra.

The accuracy of corrections appears to depend on the method used for contamination removal. The Python package molecfit is an alternative to the typical standard star method (Ulmer-Moll et al., 2019 [19]). Using molecfit, synthetic models of the Earth’s atmosphere are created based on observation conditions. Implementing the molecfit telluric model corrections yields significant variations in correction quality attributable to atmospheric parameters throughout the observing period.

This study investigates how accurately the molecfit telluric correction method attenuates flux due to telluric contamination to noise levels. Through statistical assessments evaluating independent molecular contaminating species and varying telluric absorption depths, this paper endeavors to provide insights into the accuracy of telluric corrections. Thus, conclusions can be made to highlight the importance of telluric model improvement when observing for Earth-like exoplanet transmission spectra, inform future observation proposals, and expand what is considered retrievable data from corrections.

2 Observations

The observations were conducted using the GRACES high-resolution spectrograph on the Gemini North telescope. These observations utilized the transit method (Section 4) to probe the atmosphere of WASP-33b, an ultra-hot Jupiter (UHJ) exoplanet.

The files from the observations are sorted by time and then corrected to the Barycentric Julian Date (BJD), enabling precise time-scale comparisons. The BJD is calculated using the Modified Julian Date (MJD), retrieved from the observation FITS file headers. Adding 2,400,000.5 to the MJD converts this value to the BJD (The Ohio State University [17]). For each file, the BJD of the first file can be subtracted from that of each subsequent file and then multiplied by 24 to convert to hours. This is done to improve comprehension and is utilized in all relevant figures throughout this paper.
3 Target Selection

WASP-33b is classified as a UHJ, a subclass of exoplanets within the hot Jupiter class. UHJs are distinguished by their extreme temperatures which tend to cause inflated atmospheres with chemically exotic compositions. These planets typically have equilibrium temperatures greater than 2,000 Kelvin. WASP-33b, in particular, is interesting as it is not only one of the hottest UHJs ever discovered, but is also on a very misaligned orbit which can offer fascinating insights into planetary formation processes (A. Collier Cameron et al., 2010 [2]; L. Finnerty et al., 2023 [7]; B. Gaudi et al., 2017 [8]).

In addition to misaligned orbits, UHJ exoplanets are characterized by their close orbital proximity to their host stars and short orbital periods. WASP-33b has a 1.22-day orbital period (Von Essen et al., 2014 [21]). Further, WASP-33b is orbiting a rapidly rotating A-star which contributes few stellar absorption lines to the observed spectra (A. Slettebak et al., 1980 [14]). This host star, HD 15082 or WASP-33, is a Delta Scuti variable, a subclass of variable stars otherwise known as dwarf cepheids. WASP-33 is additionally significant as it is the first Delta Scuti variable known to have a planet.

4 Methods

4.1 Transit Method

The first detection of an exoplanet using the Transit Method was shortly followed by the first atmospheric characterization of an exoplanet using this method (Charbonneau et al., 2000 [4]; Charbonneau et al., 2002 [3]). Since then, the transit method has emerged as an invaluable tool in the field of exoplanetary research. Compared to other methods, such as radial velocity or gravitational microlensing, the transit method offers unique advantages, notably in its ability to directly observe exoplanetary transits to glean detailed information about various physical characteristics. The transit method is adept at determining an exoplanet’s size, orbital period, orbital inclination, planetary density, atmospheric composition and dynamics, albedo, temperature, transit timing variations (TTVs), transit duration variations (TDVs), and confirming the existence of an exoplanet.

Conducting exoplanet observations using this method involves monitoring the light curve of a star and identifying periodic dimming caused by an orbiting planet passing between the host star and the Earth (Figure 1). When this transit occurs, a portion of the light from the star passes through the atmosphere of the exoplanet, resulting in absorption from the molecules of the atmosphere. This molecular absorption appears in the spectra and can be utilized to determine atmospheric composition.

Combining their large transit depths, short transit durations, frequent transits, and high contrast data, UHJs are ideal targets for the transit method, taking advantage of the method’s bias towards these types of exoplanets. Due to their large size and proximity to their host stars, these exoplanets produce pronounced and frequent transits, making them easily detectable. Furthermore, their high temperatures result in significant atmospheric expansion during transit, facilitating detailed spectroscopic analysis.

Telluric contamination with precise corrections are particularly relevant when searching for biosignatures, especially as it pertains to molecular evidence of life in exoplanetary atmospheres. While not all proposed biosignatures are found in abundance in the Earth’s atmosphere, many chemical markers of life are—such as O2, CH4, CO2, and more (Thompson et al., 2022 [18]; Schwieterman et al., 2018 [11]). Molecular species of high abundance in the Earth’s atmosphere contribute greater contamination, effectively burying the potential biosignature signals in the contaminated regions.
Figure 1: Transit exoplanet detection method demonstrating decreases in the total light received from the host star due to the exoplanet orbital path. Absorption due to the molecules comprising the atmosphere of the exoplanet is visible during transit. (Cowley et al., 2014 [6]).

4.2 Telluric Corrections

Telluric corrections can be performed in several ways, and have historically been conducted using the standard-star method (Vacca et al., 2003 [20]). The standard-star method relies on the observation of a nearby star with known spectral features similar to those of the target star. By comparing the spectra of the stars, telluric absorption features can be identified and removed. However, this method is not without limitations. The standard star method is sensitive to variations in atmospheric conditions and relies on the availability of suitable standard stars, as well as the additional observation capacities and facilities necessary. Primarily, the sensitivity to variations in atmospheric conditions introduces difficult to account for errors. Therefore, this study examines molecfit corrections as a potential alternative that requires less observation time and does not necessitate a suitable similar star nearby.

Corrections performed with the molecfit package utilize synthetic models of the Earth’s atmosphere to remove telluric contamination from observations (Figure 2). In order to understand the sensitivity of molecfit corrections to atmospheric variations, statistical methods (Section 4.3) were employed.

The molecfit synthetic model is created using line-by-line radiative transfer (LBLRT), generating a model of the Earth’s atmosphere based on contaminating species and atmospheric conditions (Smette et al., 2015 [15]). This model is then normalized such that no telluric contamination is equal to 1. The original flux data from observations, which includes contributions from the Earth’s atmosphere, stellar lines, and the exoplanet atmosphere, is normalized to 1 and then plotted. Then, the models can be divided from each observation file, ultimately producing the corrected flux spectra.

Peaks in the spectrum corresponding to individual telluric absorption lines were identified in the lowest airmass file. The peak locations were then used for all subsequent files to remain consistent in placement. Slight movements and fluctuations in the deepest wavelength location of each peak, as well as the width of the peaks, were accounted for by incorporating a window of data points around the peaks in the models. These windows were then translated to the corrected flux files. These windows are used to compare the corrections on locations where there are peaks, and therefore contamination, separately from the corrections on the entire spectrum. This removes any contribution to the statistical assessments from the continuum at 1 that may have altered results.

The atmospheric parameters and an observation condition were retrieved from the FITS file headers. The parameters examined include airmass (Figure 9), temperature,
humidity, and pressure. The observation condition analyzed is the signal to noise ratio (SNR). These parameters can be plotted using the observation files as they change over time (Figure 3). The time is scaled from the BJD (Section 2) to demonstrate the hours of the observing period.

As the statistical assessments were performed on the corrections, shape similarities (constant, linear, or parabolic) of the results were compared to the observation graphs (Figure 3) to examine dependency on specific atmospheric parameters or the observation condition.

## 4.3 Statistical Methods

### 4.3.1 $\chi^2$

The first implemented statistical method is the Reduced $\chi^2$ ($\chi^2_r$) test. This test is commonly used to compare observed data distribution to a model to determine similarities. This approach is therefore superior to other statistical tests like R-squared which does not accurately measure the goodness of fit for disproportionately distributed data (Colton, 2002). For this data, this approach also outperforms the Shapiro-Wilk and Anderson-Darling tests which test for normal distributions and not distributions according to a model.

The $\chi^2_r$ test is performed as shown in Equation 1. In order to perform the $\chi^2_r$ assessment, first the $\chi^2$ values must be obtained. The $\chi^2$ assessment is the sum of the squared differences between the corrected flux ($F$) and the continuum ($C$) divided by the variance squared ($\sigma_i^2$). Using that, $\chi^2_r$ is calculated by dividing the $\chi^2$ value by the degrees of freedom. The propagation of uncertainty is performed using Equation 2.
Figure 3: Atmospheric parameters (airmass, humidity, pressure, temperature) and an observation condition (SNR) plotted against the scaled Barycentric Julian Date. This scaling represents the time, in hours, of the observation period.
Figure 4: The results from the $\chi^2_\nu$ test (left) compared to the airmass observation data (right). The statistical assessment is performed over all wavelength regions and telluric depths. $\chi^2_\nu$ values normalized based on lowest airmass file for plots. Ideal corrections = 1.

$$\chi^2 = \sum_i \frac{(F_i - C_i)^2}{\sigma_i^2}$$

$$\chi^2_\nu = \frac{\chi^2}{\nu}$$

Equation 1. $\chi^2_\nu$ Test

The sum of the squared differences between the corrected flux ($F$) and the continuum ($C$) divided by the variance ($\sigma_i^2$)

$$\sigma_x = \sqrt{\left(\frac{\partial x}{\partial f}\right)^2 \sigma_f^2 + \left(\frac{\partial x}{\partial g}\right)^2 \sigma_g^2 + \left(\frac{\partial x}{\partial h}\right)^2 \sigma_h^2 + \cdots}$$

$$\sigma_x = \sqrt{\left(\frac{1}{f}\right)^2 * \sigma_f^2}$$

Equation 2. Propagation of Uncertainty

The mean $\chi^2_\nu$ value across all wavelength regions and telluric depths, excluding the continuum, is 1.76. This value denotes a reasonable fit quality, but this specific value is not the primary intention of this test. Instead, the test is intended to compare the quality of the fit over the course of the observations to determine the atmospheric parameter and observation condition dependency. Therefore, the atmospheric parameters and observation condition as well as the $\chi^2_\nu$ values were plotted against time to demonstrate their variations and compare the influence of each parameter.

To determine the effects of each parameter, the $\chi^2_\nu$ values plotted against time were normalized to compare to each of the parameters plotted against time. This comparison made the influence of airmass on the overall parabolic shape of the graph abundantly evident (Figure 4).

Because the data utilizes the lowest airmass file as the reference point, the assessments accurately portray ideal corrections at that airmass (Figure 4). For the $\chi^2_\nu$ test, ideal corrections are equal to 1. This makes the correction accuracy relative to that airmass as the test is a comparative statistical assessment.
Figure 5: The results from the RSS test (left) compared to the airmass observation data (right). The statistical assessment is performed over all wavelength regions and telluric depths. RSS values normalized based on lowest airmass file for plots. Ideal corrections = 0.

4.3.2 Residual Sum of Squares

The second statistical assessment method is the Residual Sum of Squares test. This test measures the discrepancy between the expected values and obtained data using the observed data (y) and predicted data (x). Additionally, the function that represents the predicted data (f(x)) here is a single value because the predicted data is constant at the continuum. This method is outlined in Equation 3.

$$RSS = \sum_{i=1}^{n} (y_i - f(x_i))^2$$

Equation 3. Residual Sum of Squares (RSS) Test

The sum of the squared difference between the observed (y) and predicted (x) data, where the function that represents x (f(x)) is the continuum.

Similar to the analysis of the $\chi^2_\nu$ test results, the data is analyzed based on atmospheric parameters and observation condition correlations as opposed to specific values. When plotted against time, it is once again evident that the primary influencing atmospheric parameter is the airmass.

Because the data utilizes the lowest airmass file as the reference point, the assessments accurately portray ideal corrections at that airmass (Figure 5). For the RSS test, ideal corrections are equal to 0. This makes the correction accuracy relative to that airmass, as the test is a comparative statistical assessment. To increase legibility, some plots utilize a slightly altered method where indicated. This method involves using the numpy average function to weight the error while averaging the corrected flux files. The corrected flux is then divided by this average to produce a somewhat normalized RSS value. In figure captions, this is indicated by the term 'Divided RSS'.

4.4 Wavelength Regions

These observations were conducted in visible and near-infrared regions of the spectrum. Within this region, there are two molecular species that contribute to telluric contamination: H$_2$O and O$_2$.

The contamination from each molecular species occurs in distinct and unique regions of the spectrum. The H$_2$O band is determined to encompass 5 regions of contamination, corresponding to wavelengths from 5875-6000Å, 6450-6600Å, 6970-7440Å, 7840-8600Å,
Figure 6: The 5 wavelength regions corresponding to H$_2$O contamination (left, cyan) and 3 wavelength regions corresponding to O$_2$ contamination (right, cyan). Peaks are identified across the spectrum in the telluric model (red) and then constrained when identifying the windows (cyan). The observed spectrum (orange) and corrected spectrum (blue) are plotted.

Figure 7: Comparison of the RSS assessments for contamination from shallow telluric absorption lines (0.01 to 50%) and deep telluric absorption lines (50 to 99.9%).

and 8800-10350Å. The O$_2$ band is comprised of 3 regions, corresponding to the wavelength regions from 6275-6350Å, 6865-6927Å, and 7585-7750Å. Each of the regions are highlighted in cyan (Figure 6), demonstrating all of the contaminated regions within the scope of this research.

While there is not universal agreement on what molecules are biosignatures, H$_2$O and O$_2$ are largely considered strong indicators of potential life. This is particularly true when they are in combination with other molecules such as ozone (O$_3$) or methane (CH$_4$).

4.5 Telluric Depths

The telluric depths correspond to the extent of contamination as given by the percentage of absorption present for each telluric line. For example, the deepest telluric line, where there is complete telluric contamination, is depicted as 100% absorption. Conversely, an area with no telluric absorption, would have 0% absorption. However, telluric lines do not appear independent or distinct from other contributing signals such as the exoplanet atmosphere or stellar lines. For example, a region with 20% absorption may have no contributing telluric contamination, as it may all be from the exoplanet signal or other factors. This makes the molecfit model of the Earth’s atmosphere, as a predictive tool for the placement and absorption depths of telluric contamination, useful.

A complicating factor in analyzing the accuracy of corrections performed at various depths is the number of peaks. Because the shallow and deep telluric contamination have
similar airmass dependency, the airmass may be more dependent upon peak quantity as opposed to depths (Figure 7). Additionally, contamination is most common in telluric depths below 40% absorption, which makes it more difficult to generalize the accuracy of corrections above 40% absorption as there are fewer data points (Figure 8). This disparity, while inherent to this type of data, must be acknowledged. The number of peaks for certain region and depth combinations when measuring the tellurics by 5% depth intervals is sometimes 0, rendering it unusable for assessment. These values are revisited in greater detail for each wavelength region (Figures 8, 13, and 17).

Because there are fewer data points, analysis of the deeper regions of contamination will be less generally applicable than the results for smaller absorption percentages.

5 Results

5.1 Ideal Airmass

Due to the significance of airmass on correction quality, it is important to consider and constrain ideal airmasses for observation proposals requiring high quality corrections. Using the statistical assessments covering all wavelengths and depths, the ideal airmass can be discerned from the corresponding values. As the RSS and \( \chi^2 \) assessments performed similarly, the RSS results will be shown for ease of comprehension. The best corrections are performed at airmasses at or below 1.07. Reasonable corrections were able to be performed up to an airmass of 1.10.

This airmass can be used to calculate the zenith angle ideal for observations to be conducted within. The calculation is performed using a simple geometric relationship between the airmass and the inverse of the cosine of the zenith angle (Equation 4).

\[
AM = \frac{1}{\cos(\theta)}
\]

Equation 4. Relationship between airmass (AM) and zenith angle (\( \theta \)).

Using this equation, the zenith angle of the ideal airmass, 1.07, is found to be approximately 20.84 degrees. The reasonable airmass, 1.10, has a zenith angle corresponding to approximately 24.62 degrees.

The regions of ideal corrections, with RSS values nearest to 0, were used to isolate the regions of ideal and reasonable corrections. Then, these regions are transposed to the
Figure 9: Airmass calculation method. All airmass values are found using Equation 4, where an airmass of 1 is found directly at zenith \( AM = \frac{1}{\cos(\theta)} = 1 \).

...times in which these corrections were made, to obtain the corresponding airmass for both regions (Figure 10).

Therefore, future observation proposals for individual resolution spectroscopy on 30m-class telescopes within the visible and near-infrared regions of the spectrum should prioritize airmass under 1.07 obtainable within a 20.84 degree zenith angle.

5.2 Telluric Depths by Wavelength Region

When the statistical assessments were performed on all regions of contamination corresponding to either molecular species, there appeared to be many similarities. It is evident that, when analyzing both H\(_2\)O and O\(_2\) across all corresponding regions and depths, there were similarities in their airmass dependency (Figure 11).

However, they are not completely identical and therefore further analysis is conducted. First, the contaminating species are separated by absorption depth.

Telluric contamination can have varying extents at any given point, where some absorption may be stronger or weaker (Section 4.5). In order to analyze depth in both contaminating species, the telluric depths are binned in intervals of 5%. This provides enough detail and accuracy to decipher patterns while generally maintaining groupings of a significant number of telluric lines.

5.3 O\(_2\) Band

The O\(_2\) band is comprised of the 6275-6350Å, 6865-6927Å, and 7585-7750Å wavelength regions. First, the regions were analyzed individually (Figure 12). The \( \chi^2_d \) values
Figure 10: Ideal corrections at airmasses below 1.07 (red) and reasonable corrections at airmasses up to 1.10 (blue). The Divided RSS assessment (left) is used to obtain the ideal correction regions and then transposed to the airmass plot (right) to discern airmass range.

Figure 11: $\chi^2$ (top) and RSS (bottom) results plotted against the time in hours for the H$_2$O region (left) and O$_2$ region (right) for all of the contaminated regions.
became progressively more dependent on the airmass when moving to the higher wavelength regions. In fact, the smaller wavelength regions seem to not be dependent on airmass at all, but rather lose their dependence on these parameters or become dependent on the SNR. The $\chi^2$ is more dependent on the SNR than the RSS assessment because of the emphasis in the $\chi^2$ test on the contribution of error (Equations 1, 2, and 3).

The $\text{O}_2$ band had 280 total peaks distributed inhomogeneously across the spectrum. Within the $\text{O}_2$ band, the 6275-6350 Å region had 43 peaks, 6865-6927 Å region had 73 peaks, and 7585-7750 Å region had 164 peaks. The distribution of peaks by wavelength reveals that shorter wavelength regions had no deep tellurics and longer wavelength regions possessed more peaks at deeper absorption percentages (Figure 13).

To demonstrate how the quantity of telluric lines may influence the quality of corrections, each were analyzed based on the entire spectrum (Figure 14) and individual regions (Figure 15). These plots indicate that the number of peaks may somewhat influence the telluric correction quality and perhaps at certain depths and wavelengths. Specifically, the shallow telluric depths where there were also more peaks in the 6275-6350 Å resembled the quality of the corrections. However, this may be a somewhat incomplete assessment in this regard as there are depth intervals of 5% where no telluric lines and therefore the ideal correction of 0 would be reported. However, for the regions with telluric contamination present the trend appears to hold. Additionally, within the 7585 to 7750 Å region, there is a potentially interesting connection. Though the shallow contamination depths appear to lose the influence, for depths exceeding 45% the RSS appears to resemble the influence of the number of telluric lines. This region also demonstrates a higher dependency on airmass (Figure 12). Therefore, it may be the number of telluric lines or the airmass dependency that alters the quality of the corrections corresponding to these depths.

The results (Figure 12) imply that the primary dependency for the $\text{O}_2$ band is on airmass, followed by no dependency or SNR constraints. Additionally, the results in Figures 12 and 13 demonstrate that airmass is especially influential as regions become either progressively longer wavelengths, have more telluric contamination, or have generally deeper telluric contamination. The results from Figures 14 and 15 demonstrate the quantity of telluric lines potentially having an influence, though this would be a result indicating that the scale of the number of peaks required varies greatly between the $\text{O}_2$ and $\text{H}_2\text{O}$ bands (Figures 18 and 19). To discern which of these three is the most dominant factor within the $\text{O}_2$ contamination, further examination would be required.
Figure 13: Quantity of $O_2$ band telluric peaks (y-axis) in each wavelength region of $O_2$ contamination (x-axis) sorted and color-coded by their absorption depths. Binned in 5% intervals, excluding peaks at depths $<0.01\%$ and $>99.9\%$.

Figure 14: Quantity of $O_2$ band telluric peaks (orange) compared to the correction quality as discerned from the RSS assessment (blue). Absorption depths (x-axis) are binned in 5% intervals, excluding peaks at depths $<0.01\%$ and $>99.9\%$. 
Figure 15: Quantity of O\textsubscript{2} band telluric peaks (orange) compared to the correction quality as discerned from the RSS assessment (blue) separated by individual wavelength regions. Absorption depths (x-axis) are binned in 5% intervals excluding peaks at depths <0.01% and >99.9%.
Figure 16: $\chi^2$ (first and third rows) and Divided RSS (second and fourth rows) for different regions of the H$_2$O band (columns). The wavelength increases for each region from the smallest (left) to the largest (right). The highest wavelength region examined is the most distinctly airmass dependent.

5.4 H$_2$O

The H$_2$O band is comprised of the 5875-6000Å, 6450-6600Å, 6970-7440Å, 7840-8600Å, and 8800-10350Å wavelength regions. Similar to the procedure for the analysis of the O$_2$ band, the first examination pertained to separating each region within the band.

Mirroring the results from the O$_2$ band, in the H$_2$O band there is an increased dependency on airmass as the regions near the infrared portion of the spectrum. However, in order to understand what may be impacting this increased dependency, just as was done with the O$_2$ band, other factors such as the number of peaks in each region and the telluric depth variation across the regions must also be considered.

The H$_2$O band had 85,437 total peaks also distributed inhomogeneously across the spectrum. Within the H$_2$O band, the 5875-6000Å region had 4,111 peaks, 6450-6600Å region had 4,734, 6970-7440Å region had 18,905 peaks, 7840-8600Å region had 17,321 peaks, and 8800-10350Å region had 40,366 peaks. The correction quality is assessed for each region demonstrating an increasing dependency on airmass at longer wavelengths and a greater quantity of telluric peaks (Figure 16). The distribution of these peaks in each region can be broken down into its absorption depth components (Figure 17).

The full spectrum (Figure 18) and each individual region of the H$_2$O band (Figure 19) are then analyzed. When comparing the quality of corrections with the number of telluric lines, there seems to be a closer relationship between these factors in the H$_2$O
Figure 17: Quantity of H$_2$O band telluric peaks (y-axis) in each wavelength region of H$_2$O contamination (x-axis) sorted and color-coded by their absorption depths. Binned in 5% intervals, excluding peaks at depths <0.01% and >99.9%.
contamination then in the O$_2$ bands. This relationship may also be as a result of the overall quantity of H$_2$O contamination compared to the far fewer total lines in all regions within the O$_2$ contamination windows.

To demonstrate how the quantity of telluric lines may influence the quality of corrections, each were analyzed based on the entire spectrum (Figure 18) and individual regions (Figure 19). These plots indicate that the number of peaks may somewhat influence the telluric correction quality and perhaps at certain depths and wavelengths. Similarly to the O$_2$ regions, the H$_2$O regions have far fewer telluric lines in shallow absorption regions. Therefore, though the relationship appears to hold despite this, there are some regions that show ideal corrections at 0 due to no or few telluric lines being present.

The results (Figure 16) imply that the primary dependency for the H$_2$O band is on airmass, particularly within the 8800 to 10350Å region, followed by no dependency or SNR constraints in subsequent regions. Additionally, the results (Figures 16 and 17) from the $\chi^2$ assessment, demonstrate that airmass is especially influential as regions become either progressively longer in wavelength, have more telluric contamination, or have generally deeper telluric contamination. The results (Figures 18 and 19) indicate that the quantity of telluric lines may partially explain this. However, in order to make a conclusion as to which of these three is the most dominant factor within the H$_2$O contamination, further examination would be required.

6 Conclusion

Observing transiting exoplanets with ground-based telescopes and high-resolution spectrographs is essential for obtaining individual line resolution, essential for the examination of exoplanet atmospheric spectra. Ground-based observations inherently introduce spectral contamination from the Earth’s atmospheric absorption, also known as telluric contamination, which must be accurately corrected to confidently decipher molecular signatures in exoplanet atmospheres.

From high-resolution observations of WASP-33b, an Ultra Hot Jupiter exoplanet orbiting an A-type star, this paper examines how accurately synthetic models of the Earth’s atmosphere made using molecfit can correct telluric contamination. Using regions in the visible and near-infrared regions of the spectrum, H$_2$O and O$_2$ contamination can be
Figure 19: Quantity of H$_2$O band telluric peaks (orange) compared to the correction quality as discerned from the RSS assessment (blue) separated by individual wavelength regions. Absorption depths (x-axis) are binned in 5% intervals excluding peaks at depths <0.01% and >99.9%.
corrected and then statistically analyzed.

\[ \chi^2 \] and Residual Sum of Squares assessments performed similarly, both indicating similar conclusions.

The molecfit telluric corrections quality appears to be most affected by the airmass of the observations. Therefore, for optimal corrections, we recommend an airmass at or below 1.07 which corresponds to a zenith angle of approximately 20.84 degrees. This is especially pertinent when assessing for signals within regions of longer wavelengths, deeper telluric contamination, and regions with more peaks, such as the 7585-7750Å O\(_2\) band and the 8800-10350Å H\(_2\)O band.

This work has a somewhat limited scope as the conclusions are specific to the class of instrument and to molecfit corrections. However, these are promising results for further work and studies examining the accuracy of corrections across methods in order to assure that conclusions drawn from ground-based spectroscopic analysis are accurate. Emerging technologies, particularly progress in the molecfit package and advancements in instrumentation, may allow for even more precise corrections. Additionally, for corrections using this instrument and molecfit, these conclusions can help inform future observation proposals to potentially allow for more data to be deciphered within regions of contamination that are otherwise presently not considered.

This work is also somewhat limited by the observations used as only one set of observations for one particular planetary system were analyzed. Future work in this regard would expand the dataset and perform corrections across multiple exoplanet atmospheric observations to eliminate the influence of outliers and anomalies.
References

12. The Ohio State University Department of Astronomy. Explanation of Barycentric Julian Date (BJD) https://astroutils.astronomy.osu.edu/time/bjdExplanation.html.


