Toward a Self-calibrating, Empirical, Light-weight Model for Tellurics in High-resolution Spectra

Christopher Leet1, Debra A. Fischer1, and Jeff A. Valenti2

1 Yale University, 52 Hillhouse, New Haven, CT 06511, USA; christopher.leet@yale.edu
2 Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

Received 2018 September 10; revised 2019 March 2; accepted 2019 March 5; published 2019 April 22

Abstract
To discover Earth analogs around other stars, next generation spectrographs must measure radial velocity with 10 cm s\(^{-1}\) precision. Since even microtellurics can induce RV errors of up to 50 cm s\(^{-1}\), achieving 10 cm s\(^{-1}\) precision requires precise modeling of telluric absorption features. The standard approaches to telluric modeling are (a) observing a standard star and (b) using a radiative transfer code. Observing standard stars, however, takes valuable observing time away from science targets. Radiative transfer codes, meanwhile, may omit microtelluric features, which are an important contributor to the RV error budget at 10 cm s\(^{-1}\). To address these issues, we present a telluric model of the self-calibrating, empirical, light-weight linear regression telluric (SELENITE) model for high-resolution spectra. The model exploits two simple observations: (a) water tellurics grow proportionally to precipitable water vapor and therefore proportionally to each other and (b) non-water tellurics grow proportionally to airmass. Water tellurics can be identified by looking for pixels whose growth correlates with a known calibration water telluric and modeled by regression against it, and likewise non-water tellurics with airmass. The model does not require line data, water vapor measurements, or additional observations (beyond one-time calibration observations), achieves fits with a \(\chi^{2}_{red}\) of 1.17 on B stars and 2.95 on K dwarfs, and leaves residuals of 1% (B stars) and 1.1% (K dwarfs) of continuum. Fitting takes seconds on laptop PCs; SELENITE is light-weight enough to guide observing runs.

Key words: methods: data analysis – techniques: radial velocities

1. Introduction
To expand the success of exoplanet searches, next generation spectrographs are aiming for sub-meter-per-second precision in radial velocity (RV) measurements. If the 10 cm s\(^{-1}\) instrumental precision goal of the Echelle Spectrograph for Rocky Exoplanets Search and Stable Spectroscopic Observations (ESPRESSO; Pepe et al. 2013) and the Extreme Precision Spectrograph (EXPRES; Jurgenson et al. 2016) is reached, we will be able to detect small rocky planets orbiting in the habitable zones of their host stars. Such high precision requires extraordinary new fidelity in spectroscopic data: high resolution, a high signal-to-noise ratio (S/N), and greater instrumental stability. In addition to controlling instrumental errors, success requires accounting for any systematic temporal changes in the spectral line profiles, which can arise from photospheric velocities or telluric contamination (Fischer et al. 2016).

Most work on modeling telluric contamination has been tested at near-infrared wavelengths where the telluric line depths are comparable to stellar absorption lines. However, the next generation optical spectrographs aiming for 10 cm s\(^{-1}\) RV precision will be affected by time-variable microtellurics that raster across the stellar spectrum because of barycentric velocity shifts. If we do not identify pixels that are producing the small perturbations to spectral line profiles, then microtellurics may dominate the error budget for extreme precision RV programs.

2. Telluric Spectra
Atomic and molecular species in the Earth’s atmosphere interact with solar radiation and produce absorption and emission lines that are imprinted in stellar spectra obtained with ground-based spectrographs. The non-water constituents (e.g., \(\text{N}_2\), \(\text{O}_2\), Ar, Ne, He) are well-mixed, and maintain a nearly fixed element ratio throughout the troposphere, stratosphere, and mesosphere. The concentration of some non-water species (\(\text{CO}_2\), \(\text{CH}_4\), \(\text{NO}_x\)) exhibit seasonal changes or modulation from post-industrial human activities. However, these gases have stable concentrations on timescales of (at least) several days. In contrast, 99% of atmospheric water vapor is confined to the troposphere and exhibits both temporal and spatial variability that can change by more than 10% on a timescale of an hour (Blake & Shaw 2011).

Figure 1 shows the telluric spectrum with a wavelength range of 4500–6800 Å obtained with the Fourier Transform Spectrograph (FTS) from the National Optical Astronomical Observatory (NOAO; Wallace et al. 1993). The strongest telluric lines are found redward of about 6800 Å and present a particular challenge for RV measurements in the near-infrared. However, the high S/N and resolution of the FTS telluric spectrum shows that the optical spectrum is peppered with microtelluric lines with depths that are only a few percent of the continuum. Many of the lines shallower than 1% in Figure 1 will disappear when convolved with the instrumental line spread function (LSF) of high-resolution (\(R=100,000\)) echelle spectrographs. The surviving microtelluric lines are, however, very difficult to discern when superimposed on to stellar spectra. Even for stars with constant RV, the barycentric velocity of the Earth causes the telluric lines to raster across stellar absorption lines with annual amplitudes up to 30 km s\(^{-1}\), producing small, but systematic, time-variable line profile variations. Optical RV programs aiming for 10–20 cm s\(^{-1}\) precision will need to account for microtelluric lines because they introduce errors that exceed the target RV precision (Cunha et al. 2014).
3. Current Best Practices

Since telluric contamination is a serious error source for high precision spectroscopy, there is a rich literature of practices for telluric modeling. These practices fall into three categories: (a) modeling using telluric standard stars (Section 3.1), (b) modeling using radiative transfer codes (Section 3.2), and (c) modeling using a principle component analysis (Section 3.3). Finally, we discuss the literature surrounding a new challenge in telluric modeling: microtelluric modeling (Section 3.4).

3.1. Telluric Standard Stars

The classical approach to removing telluric absorption features is to observe a telluric standard star close in time and airmass to the science object (Vidal-Madjar et al. 1986; Vacca et al. 2003). The science target’s spectrum is then divided by the spectrum of the standard star. Typically, early-type stars from early A to late B are chosen as standard stars because they exhibit few and weak metal lines, and their rapid rotation helps smear out the lines that remain. The high S/N afforded by bright stars means that with high spectral resolution, even shallow telluric lines are discernible. These stars have the drawback that their strong hydrogen absorption features at the Brackett and Paschen lines blend with their tellurics (Rudolf et al. 2016). As an alternative, a solar-type star can be used as a telluric standard using high-resolution solar spectra (Maiolino et al. 1996).

Using any standard star as a telluric reference model has several well-known drawbacks. First, it takes away precious observing time from an observation’s science targets, especially when high S/N requirements are to be met (Seifahrt et al. 2010). Second, its accuracy is limited by how well the standard star’s spectrum is known. Early-type stars often display spectral features such as oxygen or carbon lines in the near-infrared. Similarly, absorption line depths of solar-type stars may deviate from the solar FTS atlas due to metal abundance or surface temperature deviations, leaving residuals from the star’s intrinsic features in the telluric model (Rudolf et al. 2016). Compounding this problem, the need to pick a star close to the science target often forces the observation of less well-known stars. Finally, for telescopes with an adaptive optics system (e.g., the Cryogenic high-resolution Infrared Echelle Spectrograph; CRIRES), the change in source brightness between the science target and the standard will affect the instrumental profile (Seifahrt et al. 2010). In practice, Ulmer-Moll et al. (2019) find that standard stars consistently underperform other telluric removal approaches.

3.2. Radiative Transfer Codes

Today, radiative transfer codes are commonly used to correct tellurics. These codes produce a synthetic atmospheric transmission spectrum which is fit to a science spectrum to model its tellurics. Popular software packages which use this approach include Transmissions Atmosphériques Personnalisées Pour l’AStromomie (TAPAS) (Bertaux et al. 2014), Molecfit (Smette et al. 2015) and Telfit (Gullikson et al. 2014), each based on theLBLRTMradiative transfer code (Clough & Iacono 1995, Clough et al. 2005). This code directly synthesizes an atmospheric spectrum from the equations of radiative transfer. The code’s inputs are (a) line parameters such as intensity or wavelength, taken from the 2016 (or 2012 or 2008) high resolution transmission line database (HITRAN) (HITRAN et al. 2009, HITRAN et al. 2013, HITRAN et al. 2017), and (b) atmospheric profile information, such as the temperature and pressure during the observation. While, in general, radiative methods correct tellurics well, they face the following two challenges. First, radiative transfer codes may omit microtelluric features, which are an important contributor to the RV error budget at 10 cm s$^{-1}$ (Bertaux et al. 2014) (see Section 3.4).

Second, radiative transfer codes often struggle to model water lines. Bertaux et al. (2014) identify some cases in TAPAS where two adjacent water lines required different amounts of water for an adequate model. This is clearly non-physical (there is only one column density of water), but the authors are uncertain why this discrepancy appears.

3.3. Principal Component Analysis

Artigau et al. (2014) investigated the use of a principal component analysis (PCA) for empirically modeling telluric lines at near-infrared wavelengths. They used observations of hot, rapidly rotating stars to build a library of telluric standards with a range of water column density and airmass. The first five principal components of the telluric absorption features were used to fit telluric lines in spectra of program stars using least squares fitting. This empirical approach self-calibrates spectra
and avoids the need for atomic line data or estimates of water column density. We believe that PCA’s empirical approach is promising. However, PCA is a very generic model, and could benefit by incorporating the well-studied physics of telluric line formation. By introducing principled physical priors, we aim to improve the sophistication of this approach.

3.4. The Challenge of Microtellurics

Most methods for modeling telluric lines have been applied to lines that are redward of 6800 Å. The telluric features at these red wavelengths are easier to identify, both because the telluric lines are deeper and the density of stellar lines is decreasing. Currently there is not a robust method for modeling microtellurics. Unfortunately, simulations by Cunha et al. (2014) show that if ignored, microtelluric contamination in the optical spectrum will introduce RV errors between 0.2 and 1.0 m s\(^{-1}\), swamping the error budget of next generation RV surveys. Cunha et al. (2014) modeled microtelluric lines in the High Accuracy Radial Velocity Planet Searcher (HARPS) optical spectra using TAPAS, an online service that simulates atmospheric transmission with input from the Ether Atmospheric Chemistry Data Centre, atomic line data from HITRAN, and an LBLRTM code (Berta et al. 2014). The atmospheric temperature and pressure model for the geographic region near La Silla is updated every six hours, and the model with the closest match in time to the observations is adopted with small empirical adjustments to water vapor column density. Based on simulations with synthetic spectra, Cunha et al. (2014) expected that the improvement in RV precision for most stars would be in the range of 10–20 cm s\(^{-1}\). Achieving RV accuracies of 10 cm s\(^{-1}\) necessitates accurate modeling of microtellurics.

4. SELENITE: A Self-calibrating Linear Regression Model

We now describe the telluric model of the self-calibrating, empirical, light-weight linear regression telluric (SELENITE) model. Since water and non-water tellurics exhibit different behavior (Hadrava 2006), SELENITE treats their lines separately. First, we describe the training data used to illustrate and evaluate SELENITE (Section 4.1). We proceed to describe the model for water tellurics (Section 4.2) and evaluate its performance on the B-star HR 3982 (Section 4.3). We then describe the model for non-water tellurics (Section 4.4) and evaluate its performance (Section 4.5), before finally combining the two halves and applying them to Alpha Centaur B, a K dwarf with significant stellar features (Section 4.6).

4.1. Training Data

The training data included 51 spectra of rapidly rotating B stars observed with the fiber-fed CHIRON spectrograph (Tokovinin et al. 2013), which is located at 1.5 m telescope at the Cerro Tololo Interamerican Observatory (CTIO). The B-type stars are ideal for this calibration because they are bright and have few spectral lines, providing high S/N spectra that are relatively easy to continuum normalize. The iodine cell that is used for Doppler measurements with CHIRON was not in the light path for any of these observations. These spectra were obtained with the narrow slit mask, which yields a spectral resolution, \( R = 140,000 \) and exposure times were set to reach a typical S/N of 100. The airmass for each observation was recorded in the FITS header; however, no information was available regarding the total precipitable water vapor (PWV) along the observation’s line of sight or other atmospheric conditions.

Figueira et al. (2010) demonstrate the long-term stability of telluric lines at the level of 10 m s\(^{-1}\) (corresponding to 0.01 of a pixel) at the La Silla Observatory using the environmentally stabilized and fiber-fed HARPS spectrograph. The CHIRON spectrograph does not have the stability of HARPS, and the spectral format can drift by a fraction of a pixel from night to night. To correct for these small drifts, the spectral orders were cross-correlated to align the telluric absorption lines.

4.2. Water Tellurics

4.2.1. The Theory of Water Tellurics

Each water vapor line has a specific absorption coefficient, \( \sigma \), which depends on fundamental atomic and molecular line data, including the log\((gf)\) value, excitation potential, and the partition function. The radiative transfer equation for the intensity of light with wavelength \( \lambda \) passing through a plane-parallel atmosphere with a single species of absorber is

\[
I_\lambda = I_{\lambda,0} e^{-\sigma_\lambda n(z)},
\]

\[
\ln I_\lambda = -\sigma_\lambda \cdot n_j \cdot z,
\]

where \( I_{\lambda,0} \) and \( I_\lambda \) are the initial and final intensity, \( \sigma_\lambda \) is the absorber’s absorption cross section at wavelength \( \lambda \), and \( n_j \) is the average number density of absorbers. The path length, \( z \), is measured in units of airmass at zenith. The column density of water vapor, PWV, is \( n_j \cdot z \). If a spectrum is normalized (\( I_{\lambda,0} = 1.0 \)) the natural logarithm of its line intensity is proportional to the average absorption cross-section and the number of absorbers. While each water line will have a unique absorption cross-section, all of the water lines in an observation will share the same PWV \( (n \cdot z) \). The depth of any two water lines is therefore linearly related: by measuring the depth of an arbitrary water line (or set of lines), we can predict the depth of every other water line in the spectrum. We refer to the water telluric used to construct the telluric spectrum as the calibration telluric, and the pixel at the core of the calibration line as the calibration pixel.

As an example, Figure 2 shows two water telluric lines from the set of training spectra. Both sets of spectra (top and bottom right panel of Figure 2) have been color coded by the intensity of the pixel at \( \lambda = 5898.16 \) Å, emphasizing the correlated line growth. In the left panel of Figure 2, the correlation between the logarithm of the pixel intensity for these two water telluric features is shown to be linear, with a Pearson correlation coefficient (PCC) of 0.99, and the fitted regression line has residuals of 0.0085, comparable to the average deviation of the continuum from unity (0.01).

We now derive the precise relationship between the depths of any two water tellurics. From the radiative transfer equation, the intensities of any pair of water lines, \( (I_{\lambda_1}, I_{\lambda_2}) \), grow proportionally to each other in log space. Since the average number density of water absorbers and the airmass is a constant at any time \( t \), the constant of proportionality between the growth of two lines, as shown in Equation (3), can be physically interpreted as the ratio between the absorption cross-section at two wavelengths: \( \sigma_{\lambda_1} / \sigma_{\lambda_2} \). We denote this constant
of proportionality as \( m^\lambda_{\text{cal}} \).

\[
\frac{\ln I_{\lambda,i} - \ln I_{\lambda,\text{cal}}}{\ln I_{\text{cal},i} - \ln I_{\text{cal},\text{cal}}} = \frac{\sigma_{\lambda_i} [n_{\lambda_i} \cdot z_{\lambda_i} - n_{\lambda_{\text{cal}}} \cdot z_{\lambda_{\text{cal}}}] + \sigma_{\lambda_{\text{cal}}} [n_{\lambda_{\text{cal}}} \cdot z_{\lambda_{\text{cal}}} - n_{\lambda_i} \cdot z_{\lambda_i}]}{\sigma_{\lambda_{\text{cal}}} [n_{\lambda_{\text{cal}}} \cdot z_{\lambda_{\text{cal}}} - n_{\lambda_i} \cdot z_{\lambda_i}]} = \frac{\sigma_{\lambda}}{\sigma_{\lambda_{\text{cal}}}} = m^\lambda_{\text{cal}}. \tag{3}
\]

A similar linear regression is carried out to empirically relate every other pixel in the spectrum to the calibration pixel, implying an equation of the form \( \ln I_{\lambda_i} = m^\lambda_{\text{cal}} \ln I_{\lambda_{\text{cal}}} + b \). During this process, the \( y \)-intercept was always found to be zero, simplifying the regression model to

\[
\ln I_{\lambda_i} = m^\lambda_{\text{cal}} \ln I_{\lambda_{\text{cal}}}. \tag{4}
\]

There are two cases where this model breaks down. First, when a telluric line saturates, it leaves the linear regime of growth and does not obey Equation (1). Fortunately, in both our water and non-water analysis, however, we find no telluric deeper than 50% of the continuum between 4500 and 6800 Å and so no saturated telluric. Saturated tellurics are therefore considered outside the scope of this paper. Second, if the instrumental LSF changes, each telluric’s profile and therefore its value for \( m^\lambda_{\text{cal}} \) changes. SELENITE does not model instrumental errors, and these variations can only be handled by observing new training data under the new LSF. Fortunately, at CHIRON’s resolution tellurics are marginally resolved, attenuating LSF changes. In practice, CHIRON’s LSF is relatively stable over years, allowing 2012 K-dwarf observations to be fit by a model built on 2014 B-star observations (Section 4.6).

The correlated growth of water tellurics can also be exploited to identify water tellurics. The PCC of each pixel’s growth with the calibration pixel can be measured, and each pixel whose PCC exceeds a threshold, \( k \), can be flagged as containing a water telluric. Usefully, SELENITE can discover new water tellurics not contained in HITRAN and correct the position of HITRAN’s water tellurics.

Three additional tests are applied to pixels with PCC > \( k \) to eliminate false positives. First, the LSF for CHIRON has an FWHM of 3 pixels. Therefore, we require a minimum of three consecutive pixels with PCC values that exceed \( k \). Single or double pixels are assumed to be spurious. Second, because telluric lines have Gaussian profiles, the cluster of flagged pixels must pass a peak detection algorithm. Finally, the high-resolution FTS solar spectrum (Figure 1) indicates that telluric lines appear in clusters rather than as single isolated lines. Any isolated telluric without another telluric within 10 Å is therefore rejected.

### 4.2.2. Establishing a PCC Threshold

The threshold PCC (\( k \)) for flagging pixels with a telluric signal must be chosen to minimize both the number of both spurious detections (false positives) and the number of missed telluric lines (false negatives). This critical step ensures that the model telluric spectrum will have the highest possible fidelity. If spurious features are included in a model, they will be used to assign zero weight pixels, resulting in lost data for the RV cross-correlation. If telluric features are missed in a model, they will remain in the stellar spectrum and increase the RV errors.

The selection process begins by profiling the false-positive rates of different values of \( k \). The correlation between a calibration pixel and a noise pixel in the data set is simulated by generating \( n = 51 \) points of the form \( [\ln(I_{\text{cal}}), \ln(I_i)] \). The values of \( \ln(I_{\text{cal}}) \) evenly fill the range \([-1, 0]\) and represent a range of possible calibration line depths, while values of \( \ln(I_i) \) are drawn at random from a Gaussian distribution with \( \sigma = 0.01 \), representing shot noise typical of the CHIRON spectra (\( S/N \sim 100 \)). The PCC for each set is recorded, and the process repeated for 100,000 trials. The results are summarized in Figure 3. For the level of simulated noise, roughly 1% of pixels yield a PCC of 0.323; 0.1% of pixels have a PCC above 0.425 and fewer than 0.01% of pixels generate a PCC > 0.506. Since single and double pixel clusters with a PCC above the threshold are rejected, assuming that each pixel’s noise is independent, a threshold of \( k = 0.425 \) has just a 0.1% of \( 10^{-7} \% \) chance of generating a false positive. Since the CHIRON spectrum has about 200,000 pixels, this threshold has just a 0.02% chance of generating a false positive.
Once a threshold PCC is established, the minimum line depth detectable under the threshold in spectra with $S/N \sim 100$ is evaluated. A PCC threshold that is too high will fail to detect shallow lines (generating false negatives), reducing the sensitivity of the model. We again generated points representing pixels from 51 spectra with the form $[\ln(I_{\text{cal}}), \ln(I_i)]$. The calibration line depth, $\ln(I_{\text{cal}})$, was again evenly distributed across the range of $[-1, 0]$, while the pixels representing $\ln(I_i)$ were scaled according to $\ln(I_i) = c \cdot \ln(I_{\text{cal}})$. By randomly selecting values of $c \in [0, 0.07]$, these points represent telluric line depths of $\leq 7\%$. Gaussian noise consistent with $S/N \sim 100$ was then added to $\ln(I_i)$, and the percentage of time that the PCC was greater than $k$ for pixel pairs was recorded. This simulation was repeated for 100,000 trials, and the results show that 90\% of lines deeper than 2.3\% and 99.9\% of lines 3\% of the continuum will be identified as significant, reducing the sensitivity of the model.

4.2.3. SELENITE’s Water Telluric Model

The steps taken to identify and model the water tellurics in Section 4.2.1 are summarized below.

1. The PCC of each pixel’s growth with a calibration pixel is calculated. A threshold PCC, $k$, is established, and pixels with PCC $> k$ are flagged as significant.
2. Single or double pixels with PCC $> k$ are rejected as spurious.
3. The training data set is coadded and a peak detection algorithm is applied to each cluster of more than three pixels. Clusters which do not contain a peak are rejected as tellurics.
4. Any cluster of flagged pixels with no other cluster with 10 Å is rejected as a telluric feature.
5. Linear regression is carried out on pixels that are flagged as tellurics to measure $m^{\lambda}_{\text{cal}}$ relative to a pre-identified calibration pixel. The wavelength, regression coefficient, PCC, and water/non-water classification of each flagged pixel is then stored in a database.

The wavelength, linear coefficient, PCC, and a flag identifying the pixel as water is stored for each pixel that has passed the selection criteria for water tellurics is stored in a database. Table 1 lists an excerpt of a database generated from the training data’s content using the 5901.6 Å telluric as a calibrator. To generate a model of telluric water lines, the intensity of the central pixel in a calibration line is measured and information in the database is used to generate water tellurics for every pixel in the spectrum:

$$\ln I_{\lambda_{i}} = \begin{cases} \frac{m^\lambda_{\text{cal}}}{\sigma_{\lambda_{5901.6}}} \cdot \ln I_{\text{cal}} & \text{when } \lambda_{i} \in \text{valid peak } \land \text{PCC}_{\lambda} > k, \\ 0 & \text{otherwise} \end{cases}$$

where $m^\lambda_{\text{cal}}$ is the ratio of the effective absorption cross-section at pixel $i$’s wavelength, $\lambda_{i}$, to the effective absorption cross-section at the calibration line’s wavelength (or the weighted average cross-section for an ensemble of calibration lines), $\lambda_{\text{cal}}$. $I_{\text{cal}}$ is the intensity at the calibration line wavelength, and $k$ is the threshold correlation coefficient indicating telluric presence. Generation of the telluric water model takes less than 3 minutes on a 2015 Macbook Air with a 2.2 GHz Intel Core i7 processor and 8 GB of 1600 MHz DDR3 RAM and allows for the

---

**Figure 3.** Left panel: a cumulative histogram for 100,000 trials to measure the PCC between a designated calibration pixel $\ln(I_{\text{cal}})$ and pixels representing $\ln(I_i)$ with only Gaussian noise scaled to $\sigma = 0.01$. Right panel: the probability of detecting a signal as a function of telluric line depth in the presence of the same level of Gaussian noise. The purple solid, red dashed, orange dashed–dotted, and gray dotted lines show the 90, 99, 99.9, and 99.99 limits, respectively.
identification of variable numbers of telluric-contaminated pixels, depending on the PWV.

This is valuable since, as Figure 2 shows, water telluric size can vary by an order of magnitude. On nights with high PWV, at a threshold $k$ of 0.425 (see Section 4.2.2), up to $\sim$4150 pixels in our training spectra were contaminated, and 3.1% of the pixels were under 6800 Å. On dry nights, as few as $\sim$1700 pixels were contaminated, and 1.2% of the pixels were under 6800 Å. This is a saving of $\sim$75% of an order.

4.2.4. Identifying and Modeling Water Microtellurics

SELENITE is successful at identifying relatively shallow telluric features. Figure 4 shows the training set spectra for the wavelength range between 5075 and 5120 Å of our training spectra were contaminated, and 3.1% of the pixels were under 6800 Å. On dry nights, as few as $\sim$1700 pixels were contaminated, and 1.2% of the pixels were under 6800 Å. This is a saving of $\sim$75% of an order.

Figure 4. Left panel: segments of 51 overlapped CHIRON spectrum in the wavelength range between 5082 and 5094 Å, color coded by the airmass. It is difficult to pick out telluric lines in this image. However, when the pixels are color coded according to the PCC (middle panel), several microtellurics can be detected with high confidence. Zooming in on the wavelength segment at 5086 Å (right panel), the correlated pixel structure for identified weak telluric lines appears to be cleanly identified.

4.3. Results for Water Tellurics

4.3.1. Model Goodness of Fit

We evaluate SELENITE’s goodness of fit using the B-star HR 3982’s telluric spectrum. The HR 3982 spectrum used was generated by averaging three unique observations taken over 40 minutes to drive up its S/N. The goodness of fit was calculated by the data reduction pipeline ($\chi^2_{\text{red}}$) test statistic. HR 3982’s observed flux was treated as the true model, $F_{\text{obs,i}}$, SELENITE’s model of the flux as the data, $F_{\text{model,i}}$, and the error calculated by the data reduction pipeline (0.75% of continuum), scaled by (a) the root of the number of spectra coadded ($\sqrt{3}$) and (b) the root of model’s flux ($\sqrt{F_{\text{model,i}}}$) as the statistical errors, $\sigma_{\text{model,i}} = 0.0075 \sqrt{\frac{F_{\text{model,i}}}{3}}$.

First, to estimate the data quality independent of telluric removal, we measured the $\chi^2_{\text{red}}$ of a 3200 pixel wavelength range unaffected by telluric lines, 4892–4952 Å, with unity. We found a $\chi^2_{\text{red}}$ of 1.03, suggesting that our errors were well-calibrated. Next, the $\chi^2_{\text{red}}$ of our model’s fit in a 3200 pixel wavelength range with heavy water tellurics, 6472–6545 Å, was measured. This range was chosen because (a) it contains the most intense water tellurics bluewards of 6800 Å and (b) it was free from stellar features. Only pixels where a telluric was...
detected were included in the $\chi^2_{\text{red}}$ calculation. A 25 pixel range from 6521.5 to 6522.5 Å was found to have errors 20 times higher than any other error, and thus this region was flagged as an outlier and excluded. The $\chi^2_{\text{red}}$ of the telluric model was found to be 1.25. In particular, the line cores were fit well, with a $\chi^2_{\text{red}}$ of 1.11. To reach a similar $\chi^2_{\text{red}}$ in the affected and unaffected region, errors in the affected region need to be increased by $\sim 10.5\%$.

Figure 6 (top panel) plots a 5 Å excerpt from the affected region, with HR 3982’s spectrum shown in purple and our model shown in blue. The fit’s residuals deviate from unity by 1.0% on average, comparable to unaffected regions of the spectrum.

4.3.2. Relative Contribution of PWV and Airmass to Water Line Depth

A further result is that the contribution PWV to water line depth generally dominates over the airmass. As an example, Figure 7 shows that a low airmass ($z = 1.144$) observation of the 5900 Å water lines can exhibit significantly greater line depth than a subsequent higher airmass ($z = 1.454$) observation because of changes in PWV. While the water column density for an observation depends on both the average number density of absorbers along the line of sight (PWV) and the path length (airmass), PWV can vary by as much as an order of magnitude while airmass generally ranges between 1 and 2. In general, water line depth only weakly correlates with airmass. This lack of correlation can be exploited to distinguish water and non-water lines.

4.4. Non-water Tellurics

In this section, telluric absorption lines from molecules other than water are considered. Like water tellurics, each non-water telluric can be modeled by the radiative transfer equation for a
plane-parallel atmosphere and thus its signal intensity given by 
\[ n_j \sigma_\lambda \cdot n_j \cdot z, \] where \( n_j \) is the number density of the molecular species, \( j \). In theory, therefore, each non-water species could be modeled by measuring the depth of a calibration line, similarly to water tellurics.

Unlike water tellurics, however, non-water tellurics have no equivalent of PWV. Ignoring small seasonal variations in gases such as CO\(_2\), \( n_j \) is approximately spatially and temporally fixed. Therefore the column density of non-water lines can be approximated as only varying with airmass; by measuring airmass, we can predict the depth of every non-water line in the spectrum. As an example, Figure 8 (right panel) shows that over our observed range of airmass (\( z \) between 1.1 and 1.8) the signal intensity of the oxygen telluric feature at 6277.7 Å (Figure 8, left panel) is well fitted by the linear regression model 
\[ \ln(I_{6277.7\,\text{Å}}) = m \cdot z + b. \] The slope of the regression model, \( m \), measures \( \sigma_\lambda \cdot n_j \). Another difference from a regression against a calibration line is that the y-intercept (a fictitious extrapolation to zero airmass) is small, but non-zero.

Like water lines, non-water lines can be identified by measuring the correlation of their growth with airmass. Each pixel whose growth’s PCC with airmass is above a threshold, \( k \), is assumed to have non-water telluric and undergoes the same procedure as water telluric pixels. Again, this potentially allows for the detection of tellurics not listed in the HITRAN database.

Non-water lines can be readily distinguished from water lines because non-water lines have a low correlation with the water calibration pixels but a high correlation with airmass, and vice versa for water lines (see Section 4.3.2). Separating components that vary with airmass from those that do not is a benefit of SELENITE that might be useful outside the scope of this paper, as in the near-infrared, where H\(_2\)O, CO\(_2\), and CH\(_4\) lines mix. When a water and non-water line blend, the composite line can have a significant correlation with both the water calibrator and airmass. A regression model is not fit to composite lines, but they are flagged in the database.

The number density of a non-water species, of course, is not precisely spatially and temporally fixed. To increase the accuracy of the non-water model, each non-water species can...
be modeled separately, using a calibration line, following the method given for water lines. As we show in Section 4.5, however, species like oxygen can be well-modeled only measuring observation airmass, achieving a significant reduction in program complexity and computation time.

4.5. Results for Non-water Tellurics

We evaluate SELENITE’s goodness of fit using the B-star HR 3982’s telluric spectrum following the procedure described in Section 4.3.1. This time, however, we measured the $\chi^2_{\text{red}}$ of the models fit from 6257 to 6328 Å, a 3200 pixel wavelength range which encompasses the heart of the 6280 Å $\mathrm{O_2 \gamma}$ atmospheric band. Only pixels where a non-water telluric was detected were measured. The $\chi^2_{\text{red}}$ of the telluric model was found to be 1.17. To reach a similar $\chi^2_{\text{red}}$ in the affected and unaffected region, errors in the affected region need to be increased by 2.0%. Figure 9 plots the model’s fit to two oxygen doublets in the $\mathrm{O_2 \gamma}$ atmospheric band of HR 3982. The fit’s residuals deviate from unity by about $\sim0.75\%$ on average, comparable to the unaffected regions of the spectrum.

Unfortunately, there are no non-water species with telluric lines other than oxygen bluewards of 6800 Å, so we cannot evaluate our model on other species. Fundamentally, however, any non-water species whose number density is roughly spatially and temporally invariant should behave as oxygen does.

4.6. Modeling Tellurics in a K-dwarf Spectrum

Late-type stars display complex absorption features. These absorption features do not complicate SELENITE’s non-water modeling, which only measures airmass, but they do complicate water modeling, since they may blend with a calibration pixel’s line. To compensate for the loss of any given calibration pixel, a large (50+) ensemble of potential calibration pixels are given in the database.

Calibration pixels that are blended with stellar lines are identified and removed as follows. Initially, a telluric model is built by regression against the average of all calibration pixel depths. If any calibration line is blended with a stellar line, the regression model will overestimate PWV and the depth of every non-blended water line, but will underestimate the depth of the blended calibration pixel’s line. This calibration pixel can then be removed from the calibration set, and the process is repeated until the calibration set stabilizes. Empirically, we find that as long as just 25% of calibration pixels remain, SELENITE generates a good fit.

We evaluate SELENITE’s fit on late-type stars with the K-dwarf $\alpha$ Centauri B. We measured the $\chi^2_{\text{red}}$ of the models fit at the 6450 Å water band described in Section 4.3.1. This measurement, however, was complicated by $\alpha$ Centauri B’s stellar lines: if a telluric line is blended with a stellar line, the model’s fit will appear incorrect. This problem was overcome by noticing that changes in the Earth’s barycentric velocity will substantially shift the stellar lines in two observations of $\alpha$ Centauri B taken months apart while leaving the telluric lines in the same position. Tellurics that are blended in the first observation will often be unblended in the second observation, and vice versa.

To illustrate, Figure 10 (top panels) shows SELENITE’s fit to two observations of $\alpha$ Centauri B, at barycentric velocities of 1860 and 20,500 $\mathrm{m \, s}^{-1}$, for the same 5 Å wavelength range shown in Section 4.3.1. In the 20,500 $\mathrm{m \, s}^{-1}$ observation, the deep line at 6475 Å seems ill fit by the model’s pair of water lines (underlined), but in the 1860 $\mathrm{m \, s}^{-1}$ observation the deep line has shifted, revealing that it was a stellar line blended with a pair of water lines which the model now fits well. The fit’s residuals, shown in the bottom panels of Figure 10, show that when tellurics are removed the two spectra are indeed the same.

When we compute $\chi^2_{\text{red}}$, if the spectrum grossly deviates from a pixel fit (by 3.0% or more of the continuum), we assume that the pixel is blended with a stellar line and reject it. Following this procedure, we found an $\chi^2_{\text{red}}$ of 2.95 and 3.17 for the 1840 and 20,500 $\mathrm{m \, s}^{-1}$ $\alpha$ Centauri B observations. This fit, while acceptable, is somewhat poorer than HR 3982’s fit, in a large part because the telluric lines often blend with the stellar line tails, disrupting their profile slightly. For example, the wings of the small telluric at 6472.5 Å (at the far left of Figure 10) are blended with a small stellar telluric, inflating the measurement of $\chi^2_{\text{red}}$.

5. Discussion

Because of the barycentric velocity of the Earth, telluric lines raster across the stellar line profiles in time-series Doppler
measurements. Even shallow microtelluric features will degrade the fidelity of high-resolution spectra and may contribute up to 0.5 m s$^{-1}$ to the RV error budget. Since the Earth induces an RV of 10 cm s$^{-1}$ in the Sun, telluric contamination is a significant challenge in the search for the analsogs of our world. In this paper, we present SELENITE, an empirical technique for identifying and modeling telluric analogs of our world. In this paper, we present SELENITE, an empirical technique for identifying and modeling telluric features in the optical (4500–6800 Å), using the following observations: (a) water tellurics grow proportionally to PWV and therefore proportionally to each other and (b) non-water tellurics grow proportionally to airmass. Water tellurics are identified by looking for pixels whose growth correlates with a known calibration water telluric and modeled by the regression against it. Non-water tellurics are identified by looking for pixels whose growth correlates with airmass and modeled by the regression against it. SELENITE has several advantages over the alternatives:

1. Run time: once the database is built (<3 minutes on a standard PC) fitting a spectrum takes several seconds, permitting SELENITE to be used at the telescope to help guide observing runs.
2. Observing time: unlike standard stars, after a one-time observation of a few dozen B stars to build the database, SELENITE requires no further observations, saving observing time.
3. Requires no atomic/molecular line data: Unlike radiative transfer codes, SELENITE does not require atomic/molecular line data. This allows SELENITE to independently verify the HITRAN line list and potentially identify missing lines.
4. Modeling Microtellurics: SELENITE can identify and model microtelluric with depths of just 2% of the continuum, which may be omitted by traditional radiative transfer codes.
5. Distinguishes tellurics that vary primarily with airmass from those that do not: although outside the scope of this paper, this feature could be very useful in the near-infrared, where H2O, CO2, and CH4 lines mix.

We acknowledge, however, that SELENITE has certain limitations. First, stellar features in the set of training B stars, (e.g., the Paschen and Brackett lines) will distort its model. This problem can be solved by interpolating over each absorption, at the cost of introducing additional uncertainty to regions of scientific interest. Second, SELENITE only varies with airmass and PWV. Other atmospheric phenomena which may affect line profiles (e.g., wind speed; Caccin et al. 1985) are not taken into account. Instrumental changes, such as a varying LSF, are also not considered, and can only be handled by rebuilding the database for each instrumental profile change. Third, SELENITE’s PCC cutoff threshold produces discontinuities. While these discontinuities are small from CHIRON’s high S/N data, at lower S/N a line’s wings may not clear the PCC threshold, truncating them.

Despite these limitations, evaluations show that SELENITE provides excellent fits. The model’s fit to regions of intense water tellurics and non-water tellurics in the B-star HR 3982 had $\chi^2_{\text{red}}$ of 1.25 and 1.17, and thus errors just 10.5% and 2.0% bigger than the continuum’s fit to unity. Further, SELENITE’s fits to the K-dwarf α Centauri B observations had $\chi^2_{\text{red}}$ of 2.95 and 3.17, despite the $\chi^2_{\text{red}}$ test statistic being inflated by stellar line blending, confirming that it provides a good fit to late-type stars. SELENITE’s average residual is 1.0% and 0.75% for HR 3982 and 1.1% for α Centauri B, comparable to the residuals of radiative transfer methods (Ulmer-Moll et al. 2019).
The authors gratefully acknowledge enabling support from the following grants NSF-1616086, NSF-MRI0923441, NASA-NNH17ZDA001N-XRP, and NASA-NNH11ZDA001N-OSS. NSO/Kitt Peak FTS data used here were produced by NSF/NOAO.

Facility: CTIO: (CHIRON).

ORCID iDs
Christopher Leet © https://orcid.org/0000-0003-2369-0481
Debra A. Fischer © https://orcid.org/0000-0003-2221-0861
Jeff A. Valenti © https://orcid.org/0000-0003-3305-6281

References
Artigau, É, Astudillo-Defru, N., Delfosse, X., et al. 2014, Proc. SPIE, 9149, 914905
Bender, C. F., Mahadevan, S., Deshpande, R., et al. 2012, ApJL, 751, L31
Bertaux, J. L., Lallement, R., Ferron, S., Boonne, C., & Bodichon, R. 2014, A&A, 564, A46
Blake, C. H., & Shaw, M. M. 2011, PASP, 123, 1302
Caccin, B., Cavallini, F., Ceppatelli, G., Righini, A., & Sambuco, A. M. 1985, A&A, 149, 357
Clough, S. A., & Iacono, M. J. 1995, JGR, 100, 16519
Clough, S. A., Shephard, M. W., Mlawer, E. J., et al. 2005, JQSRT, 91, 233
Cunha, D., Santos, N. C., Figueira, P., et al. 2014, A&A, 568, A35
Figueira, P., Pepe, F., Lovis, C., & Mayor, M. 2010, A&A, 515, A106
Fischer, D. A., Anglada-Escude, G., Arriagada, P., et al. 2016, PASP, 128, 066001
Gordon, I. E., Rothman, L. S., Hill, C., et al. 2017, JQSRT, 203
Gullikson, K., Dodson-Robinson, S., & Kraus, A. 2014, AJ, 148, 53
Hadrava, P. 2006, A&A, 448, 1149
Jurgenson, C., Fischer, D., McCracken, T., et al. 2016, Proc. SPIE, 9908, 990806
Maiolino, R., Rieke, G. H., & Rieke, M. J. 1996, AJ, 111, 537
Pepe, F., Cristiani, S., Rebolo, R., et al. 2013, Msngr, 153, 6
Rothman, L. S., Gordon, I. E., Babikov, Y., et al. 2013, JQSRT, 130, 4
Rothman, L. S., Gordon, I. E., Barbe, A., et al. 2009, JQSRT, 110, 533
Rudolf, N., Günther, H. M., Schneider, P. C., & Schmitt, J. H. M. M. 2016, A&A, 585, A113
Seifahrt, A., Käufl, H. U., Züblin, G., et al. 2010, A&A, 524, A11
Smette, A., Sana, H., Noll, S., et al. 2015, A&A, 576, A77
Tokovinin, A., Fischer, D. A., Bonati, M., et al. 2013, PASP, 125, 1336
Ulmer-Moll, S., Figueira, P., Neal, J. J., Santos, N. C., & Bonnefoy, M. 2019, A&A, 621, A79
Vacca, W. D., Cushing, M. C., & Rayner, J. T. 2003, PASP, 115, 389
Vidal-Madjar, A., Ferlet, R., Gry, C., & Lallement, R. 1986, A&A, 155, 407
Wallace, L., Livingston, W. C., & Hinkle, K. 1993, ApJS, 195, 6