The Effects of Telluric Contamination in Iodine-calibrated Precise Radial Velocities[*]

Sharon Xuesong Wang (王雪松)[1] , Jason T. Wright[2,3] , Chad Bender[4] , Andrew W. Howard[5] , Howard Isaacson[6] ,
Mark Veyette[7] , and Philip S. Muirhead[7]

[1] The Observatories of the Carnegie Institution of Washington, 813 Santa Barbara Street, Pasadena, CA 91101, USA; sharonw@carnegiescience.edu
[2] Department of Astronomy and Astrophysics, 525 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA
[3] Center for Exoplanets and Habitable Worlds, 525 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA
[4] Steward Observatory, University of Arizona, Tucson, AZ 85721, USA
[5] Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA
[6] Department of Astronomy, University of California, Berkeley, CA 94720, USA
[7] Department of Astronomy & Institute for Astrophysical Research, Boston University, 725 Commonwealth Avenue, Boston, MA 02215, USA

Abstract

We characterized the effects of telluric absorption lines on the radial velocity (RV) precision of stellar spectra taken through an iodine cell. To isolate the effects induced by telluric contamination from other stellar, instrumental, or numerical systematic RV noise, we extracted RVs from simulated iodine-calibrated spectra of three RV standard stars regularly observed by Keck/HIRES. We add in water absorption lines according to measured precipitable water vapor (PWV) contents over a one-year period. We conclude that telluric contamination introduces additional RV noise and spurious periodic signals at the level of $10–20$ cm s^{-1}, consistent with similar previous studies. Our findings show that forward modeling the telluric lines effectively recovers the RV precision and accuracy, with no prior knowledge of the PWV needed. Such a recovery is less effective when the water absorption lines are relatively deep in the stellar template used in the forward modeling. Overall, telluric contamination plays an insignificant role in typical iodine-calibrated RV programs aiming at $\sim1–2$ m s^{-1}, but we recommend adding modeling of telluric lines and taking stellar template observations on nights with low humidity for programs aiming to achieve a precision of better than 1 m s^{-1}.

Key words: atmospheric effects – techniques: radial velocities – techniques: spectroscopic

1. Introduction

The first exoplanets around main-sequence stars were discovered with the radial velocity (RV) method, where precise Doppler spectroscopy measures the wavelength shift of the host stars induced by the gravitational pull of the planets (Campbell et al. 1988; Latham et al. 1989; Hatzes & Cochran 1993; Mayor & Queloz 1995; Butler & Marcy 1996). Since then, the RV method has been used to discover hundreds of planetary systems (see exoplanets.org; Han et al. 2014) and contributed to numerous confirmations and characterizations of exoplanets discovered with the transit method (e.g., for *Kepler* follow-up observations; Marcy et al. 2014).

The current best RV precision being readily achieved is around 1 m s^{-1} (Fischer et al. 2016), and occasionally below 1 m s^{-1} for quiet and bright stars (e.g., Pepe et al. 2011), mostly commonly attainable via two wavelength calibration methods in the optical band: calibration of ThAr lamp emission lines (e.g., ELODIE and HARPS; Baranne et al. 1996; Mayor et al. 2003; $\sim400–690$ nm) and calibration of iodine cell absorption lines (e.g., Keck/HIRES and Magellan/PFS; Vogt et al. 1994; Butler et al. 1996; Crane et al. 2010; $\sim500–620$ nm). Over 20 new precise RV spectrographs are being built and commissioned as of 2018, and most of them aim at an RV precision of better than 1 m s^{-1} or even 10 cm s^{-1} (Wright & Robertson 2017). The major obstacles to achieving a higher RV precision include stellar activity-induced RV signals, instrumental effects, telluric contamination, and limitation in data analysis (Fischer et al. 2016).

Traditionally, telluric contamination has not been considered as problematic for precise RVs in the optical. It is certainly, however, a source of severe spectral contamination and a bottleneck for achieving higher RV precision in the near infrared (NIR) region (e.g., Bean et al. 2010), where a large number of deep water and methane lines reside. However, only a small wavelength range exists in the optical that has deep telluric lines, and typically, such regions are simply omitted for the purpose of precise RV analysis, either by giving them zero weights in the cross-correlation masks (for ThAr-calibrated spectra, e.g., Pepe et al. 2002) or by flagging them as bad pixels (for iodine-calibrated spectra, e.g., for Keck/HIRES).

The works by Artigau et al. (2014) and Cunha et al. (2014) have characterized and mitigated the effects of telluric contamination in the precise RV data taken by the ThAr-calibrated HAPRS. Cunha et al. (2014) focuses on the issues with "microtelluric" lines (shallow telluric absorption lines with depths $<1\%–3\%$), which are considered in the context of precise RV for the first time. Cunha et al. (2014) fit and then divided out the telluric lines in the observed spectra using synthetic telluric spectra generated by the LBLRTM package (line-by-line radiative transfer model, Clough et al. 1992; with line lists from the high-resolution transmission molecular absorption database, or HITRAN, Rothman et al. 2013) and TAPAS (Bertaux et al. 2014), which is a more user-friendly, though less flexible, package using LBLRTM. They concluded that the microtellurics have an impact (defined as the rms of the difference between RVs before and after microtelluric removal) of $\sim10–20$ cm s^{-1} for G stars observed with low to moderate airmasses, but the impact can be substantial in some cases, up to $\sim0.5–1$ m s^{-1}.

Artigau et al. (2014) uses principal component analysis (PCA) to empirically correct for telluric lines in HAPRS data (both microtellurics and the deep lines in the $\sim$630 nm region). Combining PCA with rejection masking, they have reduced the RV rms by $\sim$20 cm s^{-1} (and more significantly for the $\sim$630 nm region). More recently, Sithajan et al. (2016) characterized the effects of telluric contamination and the effectiveness of some typical remedies (masking and modeling) for emission line-calibrated spectra for the optical, broad optical (300–900 nm), and NIR. Their conclusion for the optical region is similar to the results in Artigau et al. (2014) and Cunha et al. (2014), but in the NIR, even if all of the telluric lines are modeled and subtracted to the 1% level, the residuals would still cause 0.4–1.5 m s^{-1} RV errors for M and K dwarfs. In addition, the recent work by Lisogorskyi et al. (2019) using the HARPS data on Alpha Centauri B showed that the effects of telluric contamination can be up to a few m s^{-1} in certain spectral orders, depending on the equivalent width of the water lines.

This paper characterizes and corrects for the adverse effects of telluric contamination[8] in the context of iodine-calibrated precise RVs. In particular, we focused on quantifying the RV errors and systematics induced by tellurics and the effectiveness of mitigation methods, both in idealized cases and in realistic scenarios. We used simulated spectra in order to isolate the effects of tellurics, and we chose forward modeling as the method for mitigating tellurics instead of dividing the telluric lines out (which is mathematically incorrect because convolution, i.e., line broadening by the spectrograph, is not distributive over multiplication; Muirhead et al. 2011).

The paper is structured as follows: we describe our methodology in Section 2 and detail our findings in Section 3. We comment on results with real iodine-calibrated RV observations in Section 4 and summarize our conclusions and recommendations in Section 5.

2. Methodology

To quantify the impacts of microtelluric absorption features (often referred to simply as "tellurics" below) and isolate their effects, we performed an end-to-end simulation, from the simulated iodine-calibrated precise 1D RV spectra to the data analysis process on RV standard stars. Real iodine-calibrated RV data have systematic errors stemming from errors in the deconvolved stellar template (Wang 2016) and other unknown sources (typically referred to as the instrumental RV "jitter").

We used the Keck/HIRES spectral format for our study because Keck/HIRES is one of the most widely used iodine-calibrated spectrometers with high RV precision, and it has long observing baselines on many RV standard stars. Notably, the results in this work could be of general application despite us selecting one specific instrument, because most iodine-calibrated RV instruments are similar in terms of resolution, sampling factor on the CCD, spectral grasp, and signal-to-noise ratios (S/Ns) for their observations.

Throughout this study, we used the atmospheric conditions and weather data of the Kitt Peak National Observatory as the input into our simulated spectra. This is because Maunakea is

significantly drier than most observatories, and thus the telluric lines in the iodine region (mostly water lines) are shallower. This does not alter the main conclusion of our study, which is that tellurics have a negligible impact on the RV precision for typical iodine-calibrated spectra, aiming at 1–2 m s^{-1}.

2.1. Choice of Stars and Synthetic Stellar Spectra

We performed simulations with three stars: HD 185144 (σ Dra), HD 10700 (τ Ceti), and HD 95735 (GJ 411), which are among the benchmark RV standard stars because they are bright, quiet, and not hosts of known planets.[9] They exhibit the smallest RV variation on both short timescales of days and long timescales of years. The RV data on standard stars are often good diagnostic tools for identifying RV systematics. We simulated the spectra of these three stars with the same S/N and observing cadence as the real Keck/HIRES data.

HD 185144 is a G9V star (SIMBAD), and it is among the stars most frequently observed with Keck/HIRES. It has 712 Keck/HIRES observations as of early 2016,[10] with RV rms = 2.2 m s^{-1} (Butler et al. 2017), and it has a relatively small span in barycentric velocity (often referred to as the barycentric velocity correction, or BC), [−4.7, 4.7] km s^{-1}, because of its proximity to the northern ecliptic pole. RV systematics tend to correlate strongly with BC, because the largest Doppler signal in the stellar spectrum is the BC component, and thus BC dictates both the position of the stellar spectrum on the CCD and which iodine lines it interacts with. Due to HD 185144's small BC span, its RV data sample the BC space more densely than most stars, and therefore, in general, RV systematics would show up most clearly in an RV versus BC plot of HD 185144.

HD 10700 is a G8.5V star, and it has 623 observations as of early 2016, with rms = 2.4 m s^{-1}. Its BC span is [−27.8, 26.8] km s^{-1}, one of the largest due to its proximity to the ecliptic equator. HD 95735 is an M2V star, and it has 243 Keck/HIRES observations as of early 2017. Its RV rms is 3.3 m s^{-1}, with a BC span of [−26.4, 27.0] km s^{-1}.

To construct simulated spectra, we used synthetic stellar spectra instead of observationally derived stellar spectra, because the latter contain telluric absorption lines. The synthetic spectra of HD 185144 and HD 10700 were generated using Spectroscopy Made Easy (SME; Valenti & Piskunov 1996; Valenti & Fischer 2005; the SME spectra are provided by Dr. Jason Curtis). They were generated using the best-fit stellar parameters estimated by SME using selected spectral windows that are sensitive to key stellar parameters, such as effective temperature and gravity (Valenti & Fischer 2005). The original SME output synthetic spectra are continuum-normalized and have a resolution of $R \sim$ 1,000,000, convolved with a single Gaussian spectral point-spread function. The stellar parameters for HD 185144 are $T_{\mathrm{eff}} = 5246$ K, log $g = 4.55$, and [M/H] $= -0.16$. The stellar parameters for HD 10700 are $T_{\mathrm{eff}} = 5283$ K, log $g = 4.59$, and [M/H] $= -0.36$. The best-fit $v \sin i$ values and the default

---

macro turbulence parameters for these two stars would yield spectral lines that are too broad compared to their $R \sim 300,000$ deconvolved stellar spectral templates (DSSTs, referred to as "stellar templates" below), derived observationally from Keck/ HIRES data, so we turned off rotational and macro turbulence broadening. The output SME spectra have lines that are slightly narrower than their corresponding stellar templates, but we did not convolve the SME spectra to match the line widths in stellar templates, to avoid introducing additional numerical errors. This did not affect our study since we simulated and fit observed spectra with much lower resolutions at $R \sim 70,000$ and $R \sim 120,000$.

The HD 95735 synthetic spectrum is from Coelho (2014), provided by Dr. Paula Coelho. It has a spectral resolution of $R = 300,000$ and the input stellar parameters $T_{eff} = 3600$ K, log $g = 4.5$, and $[M/H] = -0.5$, which were chosen based on visual inspection to find the best match among the synthetic model spectra grids and HD 95735's Keck stellar template, in terms of line density and line depths. These stellar parameters are also consistent with the various measurements reported on SIMBAD. At such high resolution, the synthetic spectrum does not match the stellar template in most lines due to the limitation of the molecular line data in the optical at such a high resolution.

We also generated synthetic spectra based on the BT-Settl atmospheric models (Allard et al. 2012a, 2012b; Baraffe et al. 2015) to see whether they would better match the observed spectrum. We used version 15.5 of the PHOENIX stellar atmosphere modeling code to produce synthetic spectra. The PHEONIX and Coelho models look very similar. They both have similar variance in flux with the deconvolved observed spectrum, indicating similar Doppler content in the stellar lines, which was sufficient for the purpose of our study.

2.2. Generating Synthetic Telluric Spectra

We used two software packages, TERRASPEC (Bender et al. 2012) and TAPAS (Bertaux et al. 2014), to generate synthetic telluric spectra, and verified that they provide the same answer. Both packages were based on the HITRAN line list (Rothman et al. 2013) and the LBLRTM package (Clough et al. 1992). We used the "mid-latitude summer" atmospheric profile as the input for atmospheric layers, which is appropriate for Kitt Peak. Since the telluric lines in the iodine region are mostly water and oxygen lines, we separately generated synthetic spectra for water and oxygen, without accounting for Rayleigh scattering. The output synthetic telluric spectra have a spectral resolution higher than 1,000,000. An example of a generated synthetic spectrum is shown in Figure 1.

We changed line depths of water or oxygen lines by scaling them with a power law. The water line depths varied because of changes in the amount of precipitable water vapor (PWV) in the atmosphere and because of changes in airmass, and we computed the scaled flux as $f = f_0^{PWV \times airmass}$, where f_0 is the absorption spectrum at airmass $= 1$ and PWV $= 1$ mm. This scaling worked well for water lines because they are mostly optically thin in the iodine region. As a result, the transmission follows $f = e^{-\tau}$, and the optical depth τ is proportional to the amount of water along the line of sight, i.e., PWV $\times$ airmass, in a simple plane-parallel atmospheric model.

The oxygen line depths varied mostly because of changes in airmass since oxygen is a very stable and well mixed constituent of the atmosphere. We also scaled the oxygen lines using airmass following the above formula; although the
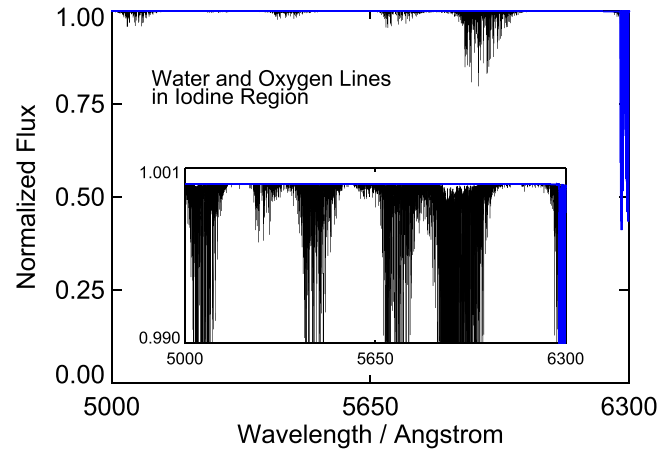

Figure 1. Telluric lines in the iodine region are mostly shallow water lines, with some moderately deep water lines near 5900 Å and very deep oxygen lines near 6300 Å. Blue lines are oxygen lines. The PWV in the plotted spectrum is 4.0 mm, typical for Kitt Peak but slightly humid for Maunakea. The atmospheric profile used is tropical (appropriate for Maunakea) with airmass 1.0. The spectrum is generated using TERRASPEC (Bender et al. 2012) at a fully resolved resolution. The insert shows the pervasiveness of microtelluric lines, i.e., lines with depths ⩽1%–3%.

oxygen lines are not all optically thin, this scaling works well for all unsaturated lines. We have verified that this simple power-law scaling is good enough to capture the line changes due to PWV and airmass by comparing the scaled spectra with models generated by TERRASPEC and TAPAS, but it is notable that modeling oxygen lines in real observations is a much more challenging task (see, e.g., Figueira et al. 2012). We used such simple scaling to vary the telluric line depths because running TERRASPEC or TAPAS is time-consuming, and it would be computationally inefficient to fold the atmospheric synthesizing code into the forward modeling code for extracting RVs from spectra.

Since our simulated spectra are based on real Keck/HIRES observations, we adopted the airmass values of these observations in our simulation to scale the water and oxygen lines. For the PWV values, we used the SuomiNet (Ware et al. 2000) water monitoring data from their station at Kitt Peak, which registers PWV values throughout the year at a cadence of 30 minutes. We used the 2016 PWV data[11] and drew randomly from the distribution of valid PWV values, but restricted it to values less than 10 mm since nights with larger PWV values are often associated with bad weather conditions that are not optimal for astronomical observations. 10 mm is essentially a conservative cut, but it does not alter the conclusion of our study. Figure 2 illustrates the PWV distribution of Kitt Peak in 2016. For comparison, the median PWV values of Maunakea and the Canary Islands are also marked, which are 2 mm and 3.5 mm, respectively. These two sites are significantly drier than Kitt Peak, having 90% or more of the nights with PWV < 10 mm (see Sarazin 2002 and García-Lorenzo et al. 2010 for Maunakea, and Castro-Almazán et al. 2016 for the Canary Islands).

2.3. Simulating Iodine-calibrated Stellar Spectra

For the three stars chosen in this study, we simulated two sets of spectra for each real Keck/HIRES observation: one set with spectral resolution $R \sim 70,000$ (70k; e.g., typical Keck/

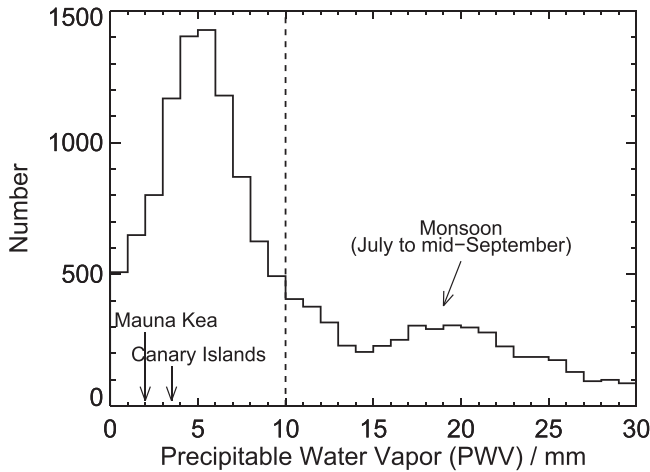[11] http://www.suominet.ucar.edu/data/staYrHr/KITTnrt_2016.plot

Figure 2. Histogram of the amount of PWV in millimeters at the Kitt Peak National Observatories from SuomiNet (Ware et al. 2000). Data were measured in half-hour intervals during the entire year of 2016. PWV values used in our simulations were randomly drawn from this distribution, but limited to PWV < 10 mm. For comparison, the median PWVs of Maunakea and the Canary Islands are also marked. See Section 2.2 for more details.

HIRES resolution), and the other set with $R \sim 120{,}000$ (120k; e.g., the resolution of Magellan/PFS2; Crane et al. 2006, 2010). The Keck/HIRES CCD has a sampling factor of ~ 3 pixels per resolution element when using the C2 or B5 decker (with $R \sim 55{,}000$, though this is normally higher and reaches up to 70,000 because of the good seeing at Maunakea). For each spectral resolution, we simulated spectra with and without photon noise. The simulated spectra with photon noise have the same S/N per pixel as the corresponding Keck/HIRES observation, typically S/N ~ 200 per pixel. All simulated spectra have one version free of telluric lines, and another one with telluric lines added, generated using the randomly drawn PWV value and the airmass value from the corresponding Keck/HIRES observation. No *intrinsic* variation in stellar velocity exists in the simulated spectra—i.e., the spectra are only Doppler-shifted according to the barycentric velocity of the Earth at the time of observation.

Below is our recipe for generating a simulated iodine-calibrated spectrum.

1. Shift the synthetic stellar spectrum to have the same barycentric velocity as the star at the epoch of the corresponding Keck/HIRES observation.
2. Generate a continuous wavelength solution for each echelle order based on the best-fit wavelength solution for the corresponding Keck/HIRES spectrum.
3. Resample the synthetic spectrum and the iodine atlas onto a wavelength grid that is four times finer than the wavelength solution grid.[12] We chose to resample at this sampling factor to match the spectral synthesis procedure in the forward modeling code. Then multiply the stellar spectrum with the iodine atlas.

4. Add telluric contamination by multiplying the spectrum produced in Step 3 with a telluric absorption spectrum scaled using the appropriate airmass and the randomly drawn PWV value.
5. Convolve each order of the multiplied spectrum with a single Gaussian line-spread function (LSF).[13] We used a Gaussian LSF with $\sigma = 1.7$ pixels for $R = 70$k and $\sigma = 1.0$ pixel for $R = 120$k. Then we resampled the convolved spectrum down to the observed wavelength pixel grid using a flux-conservative algorithm.
6. Fit for the blaze function for each order of the corresponding Keck/HIRES spectrum with a third-order polynomial using the top 1% of the flux, and add this fitted blaze function to the convolved spectrum. This ensures that the simulated spectrum has the same photon counts as the corresponding Keck/HIRES spectrum.
7. Add photon noise according to Poisson statistics, when appropriate. For the pair of simulated spectra with and without telluric lines, we added the exact same photon noise in each pixel for each pair to ensure that the change in reported RVs was purely due to the added tellurics instead of variance in the randomly generated photon noise.

RVs are extracted from iodine-calibrated spectra via forward modeling (Butler et al. 1996; see the next section), which requires a stellar template (or DSST; as mentioned in Section 2.1). We simply converted the synthetic stellar spectra into the stellar template format needed by the Doppler code (with spectral sampling and division into chunks to match the real Keck/HIRES stellar template for each star). We shifted the wavelengths of synthetic stellar spectra to be at the same wavelengths as the corresponding Keck/HIRES observations that generated the stellar templates, i.e., accounting for each star's systemic velocity and the barycentric velocity of the Earth at the epoch of observation of the stellar template. As a result, just as for real observations, all reported RVs were the stellar velocities *relative to* the star's velocity at the epoch of its template observation. We added telluric absorption lines into the stellar templates since, in reality, the stellar template is derived from on-sky observations. Eliminating telluric absorption lines from DSST is not a trivial process given the shallow line depths of water lines and the fact that they are heavily blended with the stellar lines. We have experimented with cleaning telluric lines for DSST in real observations, but, given the resolution, sampling, and S/N of the real observations, the extra uncertainties induced by this "cleaning" process introduced more errors than the errors induced by the tellurics (Wang 2016). Better algorithms for generating stellar templates may yield a different or better result, but are beyond the scope of this paper. Therefore, for our simulations, we did not use telluric-free stellar templates. We used two versions of stellar templates: PWV = 1 mm (simulating template observations taken on a very dry night) and PWV = 5 mm (taken on a typical night), both with airmass = 1.0.

2.4. The Doppler Code for Extracting RVs

We used the California Planet Survey (CPS) Doppler code to extract precise RVs from the simulated spectra. The code uses

---

[12] Ideally, one would multiply the two spectra at a higher sampling and then resample to a coarser grid. However, we chose to downsample the synthetic spectrum here before multiplication because the input stellar template being used to forward model these simulated spectra was downsampled as such to match the real Keck/HIRES templates. In this way, our synthesizing process for making simulated spectra was numerically identical to the synthesizing process used in the forward modeling code, which eliminated potential additional numerical errors.

[13] The LSF model for real Keck/HIRES observations is a multi-Gaussian model, with a central Gaussian with fixed width and satellite Gaussian pairs with fixed positions and width but varying heights (Valenti et al. 1995). We also performed simulations using a multi-Gaussian LSF input and found the same conclusion. Here, we only present results from simulations using single Gaussian LSF input for simplicity.
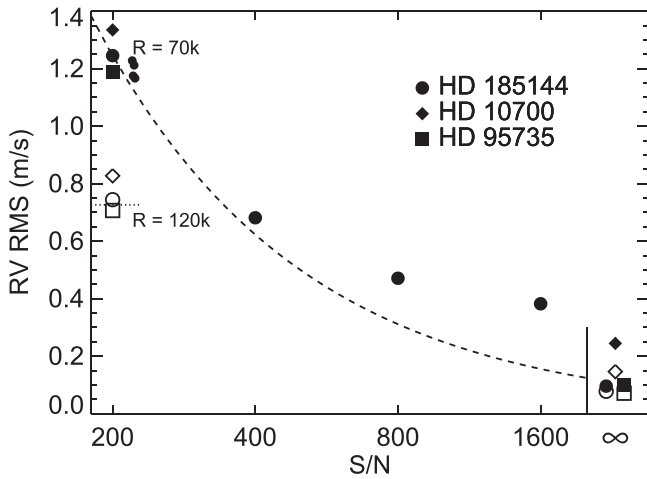
Figure 3. RV rms vs. S/N per pixel in the simulated telluric-free spectra for each star to demonstrate the precision of the adopted Doppler pipeline. Filled symbols are for the simulated spectra with resolution $R = 70,000$ or 70k, and hollow symbols are for the higher resolution spectra ($R = 120,000$ or 120k). The dashed line is the predicted RV precision scaling relation assuming rms $\propto$ S/N^{-1} (e.g., Butler et al. 1996). The dotted line marks the scaled precision at $R = 120$k assuming rms $\propto R^{-1}$. Both scaling relations are normalized at S/N = 200, rms = 1.25 m s^{-1} for HD 185144 with resolution $R = 70$k. The small dots (shifted horizontally a bit off S/N = 200 for visualization) are rms for other sets of simulated spectra of HD 185144 with S/N = 200 and $R = 70$k with different photon noise added for each set. The set of points to the right of the vertical line is RV rms from noise-free spectra (labeled with ∞ for S/N). See Section 2.4 for more information.

forward modeling to fit the iodine-calibrated stellar spectra, following the algorithm outlined in Butler et al. (1996). Section 2 of Wang (2016) has a more detailed documentation on the code, and some of its elements are described in Johnson et al. (2006, 2011) and Howard et al. (2009, 2011). Our copy of the code was kindly provided by John A. Johnson and the CPS group in 2013.

The Doppler code takes the DSST and the iodine atlas as input model spectra, and fits the observed spectra to solve for the wavelength solution and Doppler shift. Each observed spectrum at a given epoch is divided into chunks of 80 pixels (about 2 Å) and fitted independently using the Levenberg–Marquardt least-χ^2 algorithm. Later, the RVs reported from all chunks are combined through an outlier rejection and weighted averaging process to yield the final RV for the given epoch. The free parameters in our fitting for a telluric-free spectral chunk include λ_0, the wavelength for the first pixel of this chunk (Å); $\delta\lambda$, the linear wavelength dispersion (Å per pixel); the width of a single Gaussian LSF, σ (pixel); and the Doppler shift, z. Although we knew the exact width for the LSF in our simulated spectra, we chose to float σ as a free parameter because, otherwise, the reported RVs contain additional systematics, probably induced by numerical errors in the making of the simulated spectra. For the simulated noise-free spectra, we used a quadratic wavelength solution in the fitting instead of a linear one, which yields a significantly higher RV precision (20–30 cm s^{-1} versus <10 cm s^{-1}). For simulated spectra with noise added, we did not invoke a quadratic wavelength solution because it would have induced additional numerical errors for spectra with more realistic S/Ns.

Figure 3 illustrates the precision of the Doppler code for spectra with different S/N and resolution (all from telluric-free simulations). The filled black dots are rms values of the RVs for simulated observations of HD 185144 at different S/Ns with a resolution of $R = 70$k. They roughly follow rms $\propto$ S/N^{-1} (Butler et al. 1996 e.g.,), but have lower precision than predicted at high S/N because of inaccuracies in the wavelength solution for each chunk—the model used is only linear in pixel space and does not have higher-order terms. The precision is recovered once a quadratic wavelength solution for each chunk is implemented, for example, for the simulations without photon noise added, plotted in the lower right corner labeled with ∞ for S/N. The high-S/N results are not realistic; rather, they are just an estimate of the code's precision—CCDs very often enter the nonlinear regime or saturate near and beyond S/N = 400. Figure 3 shows that the code performs well since the RV rms or precision scales as expected with S/N and spectral resolution, and it is capable of producing RV precision of better than 1 m s^{-1} for spectra with high enough resolution and S/N. The RV rms values reported for the noise-free cases represent a floor of systematic RV error from the forward modeling code itself.

When fitting simulated spectra injected with telluric absorption lines, we either ignored the telluric lines (i.e., did not incorporate any special treatment for them, such as masking them out), to assess the impacts of tellurics so as to compare the results with telluric-free spectra, or incorporated a component of the telluric spectrum in the forward modeling to assess how effective modeling mitigates the effects of tellurics. When including a telluric component in the forward modeling, an additional free parameter for the PWV was introduced to fit for water absorption lines by scaling the input water spectrum in a power law with PWV $\times$ airmass (see Section 2.2). The input model water spectrum was the same one used for synthesizing the simulated spectra, with PWV = 1 mm. Because most spectral chunks have a negligible amount of water absorption, only chunks with more than 0.1% water absorption (when PWV = 1 mm) were fitted with a telluric component. We did not attempt to fit or offset the telluric lines in the stellar templates (see more in the last paragraph of Section 2.2).

When the code fit for the telluric component, it introduced an additional "pass" in the least-χ^2 fitting procedure: it first fit for the PWV in the first pass, and then fixed the PWV to the best-fit value in the subsequent passes where it fit other parameters, including the Doppler shift. The fixed best-fit PWV value was the numerical median of all fitted values from chunks with more than 0.1% water absorption.

In principle, we could have also fit the oxygen lines near 6300 Å. We chose not to do so in this paper, because the choice would not alter our conclusion. For real observations, it is extremely hard to fit the oxygen lines to a high enough precision—it is challenging for the synthetic telluric model to accurately reproduce the shapes of these deep absorption lines. The line shapes depend closely on the conditions of atmospheric layers since oxygen has a significant scale height in the atmosphere. The exact line centers or wavelengths could also change, for example due to wind (Figueira et al. 2012). As a result, oxygen line models would leave a significant residual after fitting and still affect the quality of the RVs out of these spectral chunks. In addition, the wavelength region beyond 6200 Å has significantly fewer iodine absorption lines and thus much less information content. The results of both of these effects would give rise to very large RV rms values for spectral chunks contaminated with oxygen lines, and therefore these chunks would be rejected when the code computes a final weighted average RV using RVs from all chunks.

Table 1
List of Tests and Results

Type of Obs.	Spectral Resolution	Template PWV	Fit Tellurics?	HD 185144 rms (m s^{-1})	HD 10700 rms (m s^{-1})	HD 95735 rms (m s^{-1})
Noise-free[a]						
No tellurics	70k	⋯	⋯	0.10	0.24	0.10
With tellurics	70k	1 mm	No	0.13	0.29	0.14
With tellurics	70k	1 mm	Yes	0.11	0.22	0.12
With tellurics	70k	5 mm	Yes	0.15	0.20	0.14
No tellurics	120k	⋯	⋯	0.08	0.15	0.07
With tellurics	120k	1 mm	No	0.09	0.18	0.10
With tellurics	120k	1 mm	Yes	0.08	0.14	0.08
With tellurics	120k	5 mm	Yes	0.09	0.15	0.09
S/N = 200						
No tellurics	70k	⋯	⋯	1.25	1.33	1.19
With tellurics	70k	1 mm	No	1.28	1.35	1.23
With tellurics	70k	1 mm	Yes	1.24	1.29	1.23
With tellurics	70k	5 mm	Yes	1.27	1.38	1.21
No tellurics	120k	⋯	⋯	0.74	0.83	0.71
With tellurics	120k	1 mm	No	0.75	0.85	0.75
With tellurics	120k	1 mm	Yes	0.73	0.85	0.72
With tellurics	120k	5 mm	Yes	0.75	0.89	0.73
Real Obs.						
⋯	∼60–70k	⋯	No	2.62	3.08	3.43
⋯	∼60–70k	⋯	Yes	2.71	3.13	3.34
⋯	∼60–70k	cleaned DSSTs	Yes	2.75	3.66	3.23

Note.
[a] With quadratic wavelength solution when forward modeling.

3. Results

In this section, we present the results from the simulations described above. We first quantified the effects of telluric contamination in iodine-calibrated precise RVs (Section 3.1) in terms of the amount of added RV rms (at the level of $10 \, \mathrm{cm \, s^{-1}}$), as well as the induced systematics or spurious periodic signals (with amplitudes around 10–$20 \, \mathrm{cm \, s^{-1}}$ at periods of one year and its harmonics; Section 3.2). We show that forward modeling the telluric lines effectively mitigates these effects. Then we discuss retrieving the PWV values by forward modeling the water absorption lines in Section 3.3.

3.1. Impact of Telluric Contamination on RVs: Increased RV rms

Table 1 lists the RV rms values for the three stars we adopted in various simulations (with or without tellurics, with or without fitting tellurics). Table 1 has three major sections: the top section lists results from simulations using the noise-free spectra, the middle section lists results from simulations using the S/N = 200 spectra, and the last section lists results using real Keck/HIRES observations. The first four columns specify the conditions used in synthesizing the spectra or in extracting the RVs. The first column specifies whether the synthesized spectra used have telluric absorption added. The second column contains the spectral resolution. The third column specifies which stellar template was used when extracting RVs from the synthesized spectra—a telluric-free one, or one with water absorption lines with PWV = 1 mm or 5 mm. The fourth column indicates whether a telluric component was fitted or not in the forward modeling process when extracting RVs. The last three columns list the rms values of RVs for each star in each simulation.

Comparing the RV rms values in row 1 versus row 2 and row 5 versus row 6 of the first two sections in Table 1, it is evident that telluric absorption adversely affected the RV precision, though the effect is very small. The RV rms values from telluric-free spectra versus telluric-contaminated spectra only differed by a few $\mathrm{cm \, s^{-1}}$. More precisely, without any treatment, telluric absorption added about $10 \, \mathrm{cm \, s^{-1}}$ additional RV error in quadrature to the RV rms (up to $30 \, \mathrm{cm \, s^{-1}}$ for S/N = 200 cases). This result is similar to the values reported for HARPS in Cunha et al. (2014) and Artigau et al. (2014), though smaller in general because we worked with a smaller wavelength coverage with less telluric absorption than HARPS (wavelength coverage: 378–691 nm; Mayor et al. 2003). This effect of added RV rms was more evident in the noise-free simulations, and it definitely existed for situations with realistic S/Ns as well (with the added amount of RV rms consistent with the noise-free ones). However, in practice, the added RV scatter is basically negligible considering that, for real observations, the photon-limited precision is usually much larger than $10 \, \mathrm{cm \, s^{-1}}$. This amplitude of $10 \, \mathrm{cm \, s^{-1}}$ was also on par with the amount of systematic error from the forward modeling process alone, represented by the RV rms values for the noise-free telluric-free cases.

The RV rms values listed in rows 3 and 7 of the first two sections in Table 1 demonstrate the effectiveness of modeling in mitigating telluric contamination. In almost all cases, modeling improved the RV precision, reducing the added rms by 50% or more. Modeling of tellurics did not completely recover the lost precision due to errors in modeling the water lines and the existence of telluric lines in the stellar template, as well as systematics introduced into the forward modeling procedure because of the additional parameter of PWV. Again, the noise-free cases saw a more obvious improvement than the
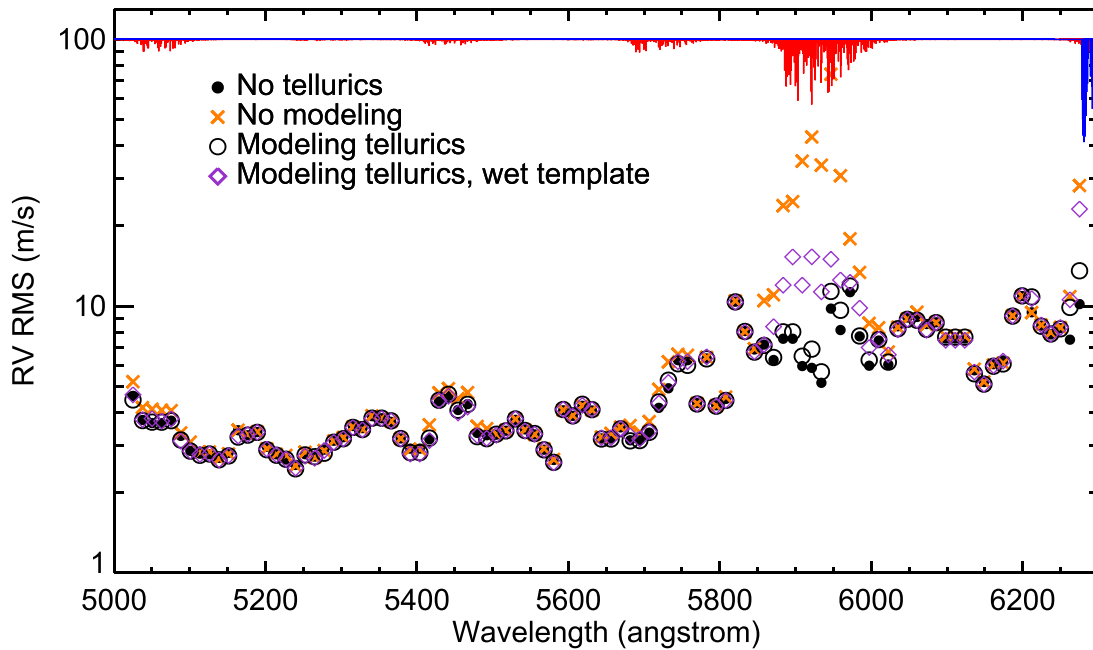
Figure 4. RV rms for each 50 Å spectral region vs. wavelength in angstrom for simulated observations of HD 185144 with resolution $R = 120k$ and $S/N = 200$. The black dots show the RV precision of telluric-free spectra, showing the photon-limited precision across the iodine region. The orange crosses show the RV precision of simulated spectra with telluric contamination injected but no treatment for the tellurics, using a variation of airmass from real observations and PWV values of the Kitt Peak Observatories in 2016. The hollow circles illustrate how the the precision improves when the telluric absorption lines are being modeled in the Doppler code (but the code uses a DSST with PWV = 1 mm). The hollow diamonds are similar to the hollow circles, but use a DSST with PWV = 5 mm. Water and oxygen absorption lines are shown on top in red and blue, with arbitrary depths.

S/N = 200 ones (in particular, HD 95735 had a slight increase in rms when modeling of tellurics was added for the $R = 70k$ and $S/N = 200$ cases, probably because the effects of tellurics/modeling were buried under noise for this resolution and S/N).

Telluric contamination in the stellar templates would certainly increase the RV rms as well. With modeling of telluric lines in the observed spectra (and without taking out the tellurics in the stellar templates), simulations using templates with PWV = 5 mm have 5–30 cm s^{-1} RV rms added in quadrature compared with the telluric-free cases. Again, these are not large compared with the typical photon-limited precision, but a general conclusion would be that stellar template observations should be taken on dry nights, if possible. For stellar template spectra with high spectral resolution and high S/N, it is plausible that the telluric lines in the stellar template could be removed by simply dividing them out, although it would not be mathematically correct, but the leftover residual would be smaller than the original water line depths. However, as mentioned in Section 2.2, we performed this experiment with Keck/HIRES templates and concluded that it does not improve the RV precision. It may require spectra with a higher fidelity (e.g., better known LSF and/or wavelength solution) for this proposed method to be effective.

Figure 4 illustrates the impact of telluric contamination in RV precision as a function of wavelength. The overall increased RV rms was due to the added RV rms from the wavelength regions with more water or oxygen absorption, e.g., around 5900 and 6300 Å (the orange crosses in Figure 4). Modeling very effectively brings down the RV rms, providing an almost full recovery from the damage (i.e., comparing the black circles versus the black dots), unless the stellar template has significant telluric absorption (PWV = 5 mm; purple

diamonds). Figure 4 is an illustrative case using simulated data on HD 185144 with S/N = 200 and R = 120k, but the general trend and conclusion were the same for simulations with other stars/scenarios.

Overall, tellurics have such a small impact on the RVs primarily because the wavelength regions with relatively heavy telluric contamination were assigned small weights when the RVs from all wavelengths were combined. The telluric-contaminated regions produced RVs with a higher rms, as shown in Figure 4, so they naturally received low weights, because weights were computed based on the RV rms of each region over time. This was particularly true for the oxygen region, where deep oxygen lines induced large RV biases as the stellar lines were red-/blueshifted back and forth across these oxygen lines due to the barycentric motion of the Earth. The entire region with oxygen absorption was actually mostly rejected and not used in the calculation of the weighted average due to the large intrinsic errors (see the difference between the hollow circles and solid dots in Figure 4). This was still true even when we added telluric modeling due to the oxygen lines in the stellar template.

Besides adding scatter to the RVs across multiple nights, telluric contamination also degraded the internal RV precision of each nightly observation. Figure 5 illustrates this effect, where the black dots show the increase in the reported internal RV errors for each observation as the PWV increases. The internal RV error for an observation was derived by evaluating the scatter in the RVs reported by all chunks internally within each observation. This error increased up to ~10% or so, consistent with the fact that the overall rms did not increase by more than a few percentage points in the telluric versus telluric-free cases (Table 1). Such an increase in the internal errors disappeared once we invoked modeling of tellurics, as illustrated by the green crosses.
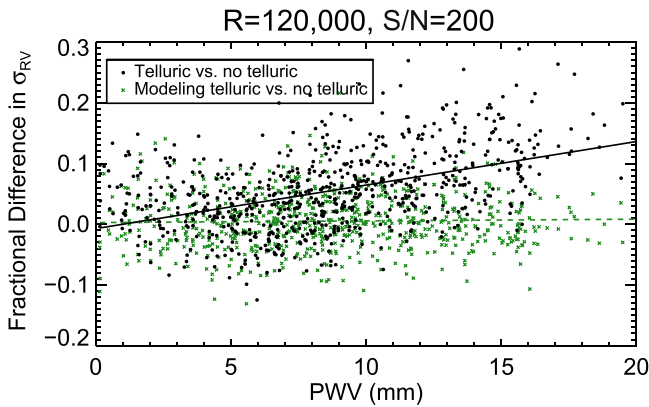
Figure 5. Changes in the internal RV error ("photon-limited" precision) for each observation due to tellurics vs. PWV. Each black dot is the fractional difference in the reported internal RV error (σ_{RV}) between a telluric-contaminated spectrum, though one not being treated for tellurics, and its paired telluric-free spectrum. The case illustrated here is for HD 185144 at $R = 120$k with S/N = 200. All green crosses are similar, each being the fractional difference between a telluric-contaminated spectrum with forward modeling of telluric lines and its paired telluric-free spectrum. The black line is a linear fit to the black dots, and the green dashed line is a linear fit to the green crosses. See Section 3.1 for more details.

There was no significant dependence on the PWVs in terms of the amount of bias induced in the RV of each observation. That is, the amount by which the RV deviated from the true RV ($0\,m\,s^{-1}$) does not appear to depend strongly on the PWVs (tested to be statistically insignificant). This lack of significant dependence on the PWVs can also be seen in Figure 6. In fact, the dependence of RV bias on PWVs only started becoming pronounced in test simulations with unrealistically high PWV values (e.g., 10 times the values used here, typically 50 mm or higher).

3.2. Impact of Telluric Contamination on RVs: Spurious Periodic Signals

Telluric contamination could also induce spurious periodic signals in RVs, typically at periods of one year and its harmonics. Such periodicity is due to the barycentric motion of the Earth, which causes the stellar lines to move back and forth on top of the static telluric lines in wavelength space. This back-and-forth motion has a period of one year, and thus the induced RV bias would also have this signature timescale. Such an "annual" systematic effect is best illustrated in RV–BC space, BC being the barycentric velocity of the star. Figure 6 illustrates this systematic error using simulations with noise-free spectra (because the effects would be buried in noise when photon noise is added). The left panels are for RVs extracted without modeling the tellurics, showing correlation or pattern in the RV–BC plane. The right panels illustrate that the RV–BC correlation is reduced with modeling of the tellurics.

Any strong correlation in RV and BC would translate into periodic signals with periods of one year and/or its harmonics. Figure 7 shows the Lomb–Scargle periodograms for our three test stars for RVs extracted from noise-free spectra without tellurics (black solid), with tellurics added but no modeling of them (orange dashed), and with tellurics and with modeling (green dashed–dotted). The black solid lines are illustrating the window function and any periodic signals induced by systematics intrinsic to the RV extraction pipeline. Comparison between the orange dashed line and the black solid lines reveals

the spurious periodic signals introduced by the telluric contamination, and they show up near 360 days, 120 days, 60 days, and so on, as expected. The green dashed–dotted lines, i.e., modeling the tellurics, basically recover the black lines, with the exception of HD 185144. For HD 185144, because of its small BC span, the telluric lines in its stellar template are never completely unblended from the telluric lines in the observed spectra being modeled, and therefore modeling of the tellurics has the least power of restoration.

As shown in Figure 6, the amplitude of the added spurious signal is of the order of $\sim 20\,cm\,s^{-1}$, which is very small given the typical photon-limited RV precision of iodine-calibrated spectra. Therefore, in typical real-life scenarios, telluric contamination does not pose a concern in terms of adding spurious signals for iodine-calibrated RV spectrographs aiming at ~ 1–$2\,m\,s^{-1}$ precision.

3.3. Retrieval of Precipitable Water Vapor via Modeling

The best-fit PWV values are a natural product of our forward modeling process. Since our code assumed no prior knowledge of PWV, how well we recovered the PWV values versus the real input values would provide a gauge on the quality of retrieved PWVs for real observations. Retrieving the PWVs in the iodine region could be very useful, because it is one of the relatively clear regions in terms of tellurics (without heavy blending of stellar lines, especially for G dwarfs, and telluric lines of multiple molecular species). For example, the PWV values derived from this wavelength region can then be used for modeling the spectrum in the NIR, where telluric lines are deeper and line blending is more severe, posing more challenges to PWV retrievals.

When using a telluric-free stellar template in the forward modeling, PWV values can be recovered to relatively high accuracy, typically to better than 0.5 mm. This is on par with or better than other PWV monitoring methods using NIR photometry (Baker et al. 2017) or the Global Positioning System (Blake & Shaw 2011). However, as mentioned in Section 2.3, it is unrealistic to have a telluric-free stellar template. As a result, the presence of telluric lines in the stellar template will bias the PWV measurements, especially when the telluric lines in the template are blended with the ones in the observed spectra being modeled. This bias is illustrated by the gray symbols in Figure 8. When a star has a small BC span, such as HD 185144 (black points), this bias affects more observations, because the relative redshift between template and observations is always small. The amount of bias depends on the depths of water lines in the stellar template (hollow symbols versus filled symbols).

We have also divided out the telluric lines in the simulated DSSTs in order to achieve better accuracy in the retrieved PWV values. This is trivial for the simulations, because we know exactly the input PWV, and the wavelength solution for the simulated DSSTs is also perfect, so we can divide out the telluric lines "perfectly."[14] As shown in Figure 8, the PWV values retrieved using a cleaned DSST match very well with the input PWV values. For simplicity, we did not list the RV

---

[14] In reality, one can never divide out telluric lines perfectly, because the observed spectrum was convolved with the spectral PSF, and mathematically the telluric absorption was multiplied on top of the stellar absorption lines before this convolution took place. Since convolution does not distribute over multiplication, dividing the telluric lines in the convolved spectrum is mathematically incorrect.
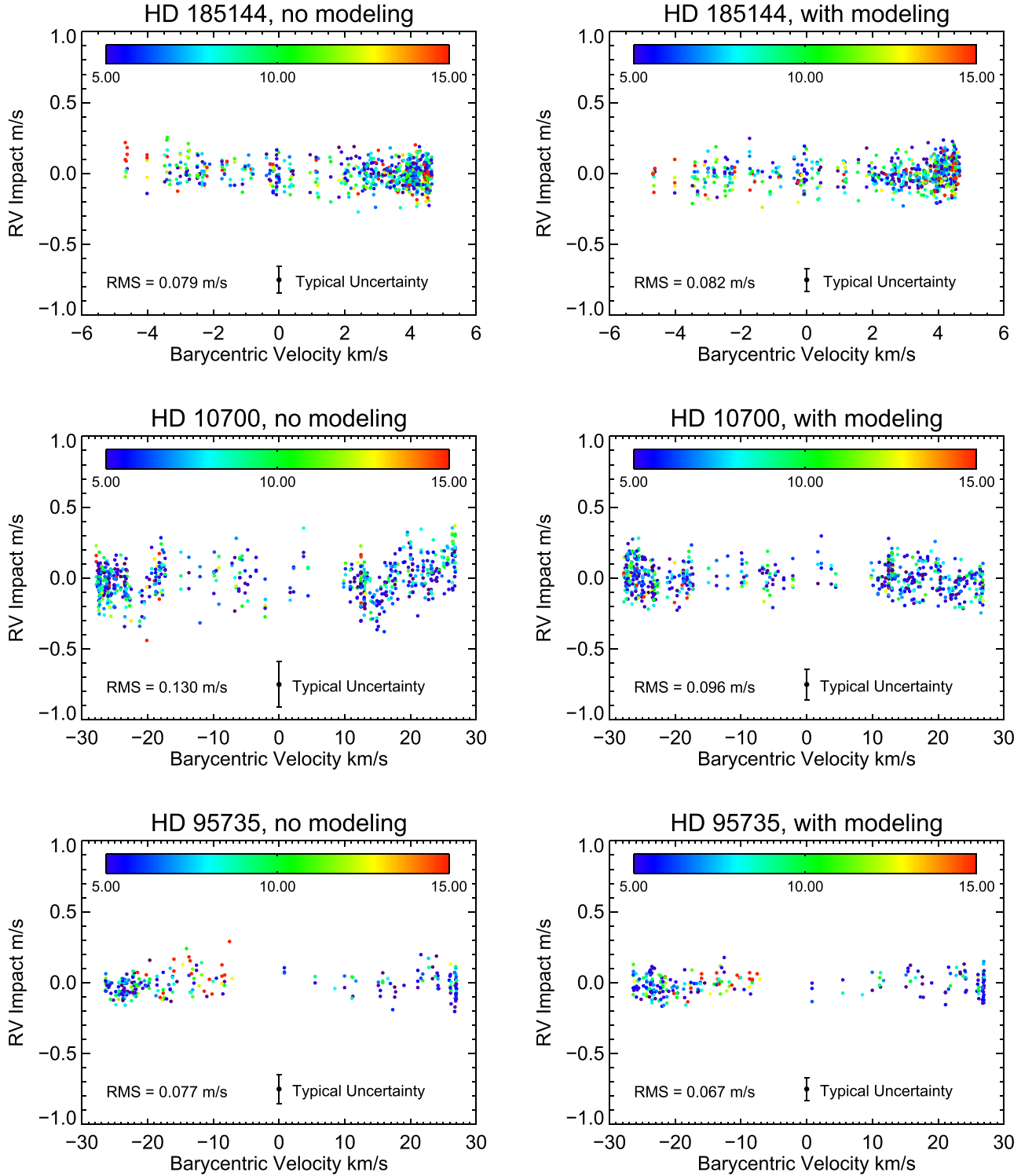
Figure 6. Difference in RVs between telluric-contaminated and telluric-free spectra vs. barycentric velocity of the star for each observation in each star (labeled in subplot titles). The left columns plot the RV difference when forward modeling of tellurics is not invoked, while the right columns show the results with modeling of tellurics. Both sets of simulations used DSST with PWV = 1 mm. Typical internal RV error for each observation is plotted and labeled as "Typical Uncertainty." The color scale denotes the PWV in millimeters. Modeling is clearly an effective method to remove systematics caused by tellurics. Note that the BC span of HD 185144 is very small. See Section 3.1 for more details.

rms values from analyses using these cleaned DSSTs, because they are essentially the same (within 1 cm s^{-1} or the numerical precision of our Doppler code) as the RV rms from the telluric-free simulations. Such a complete elimination of the effects of tellurics, however, is unfortunately not true for real observations. As listed in the last row of Table 1, RV analyses on real

observations using "cleaned DSSTs" did not return better RV rms consistently (see Section 4 for more). Therefore, it is unclear whether cleaning the telluric lines in DSSTs would result in any significant benefit besides accuracy in the retrieved PWV for a typical iodine-calibrated RV instrument with 1–2 m s^{-1} precision.
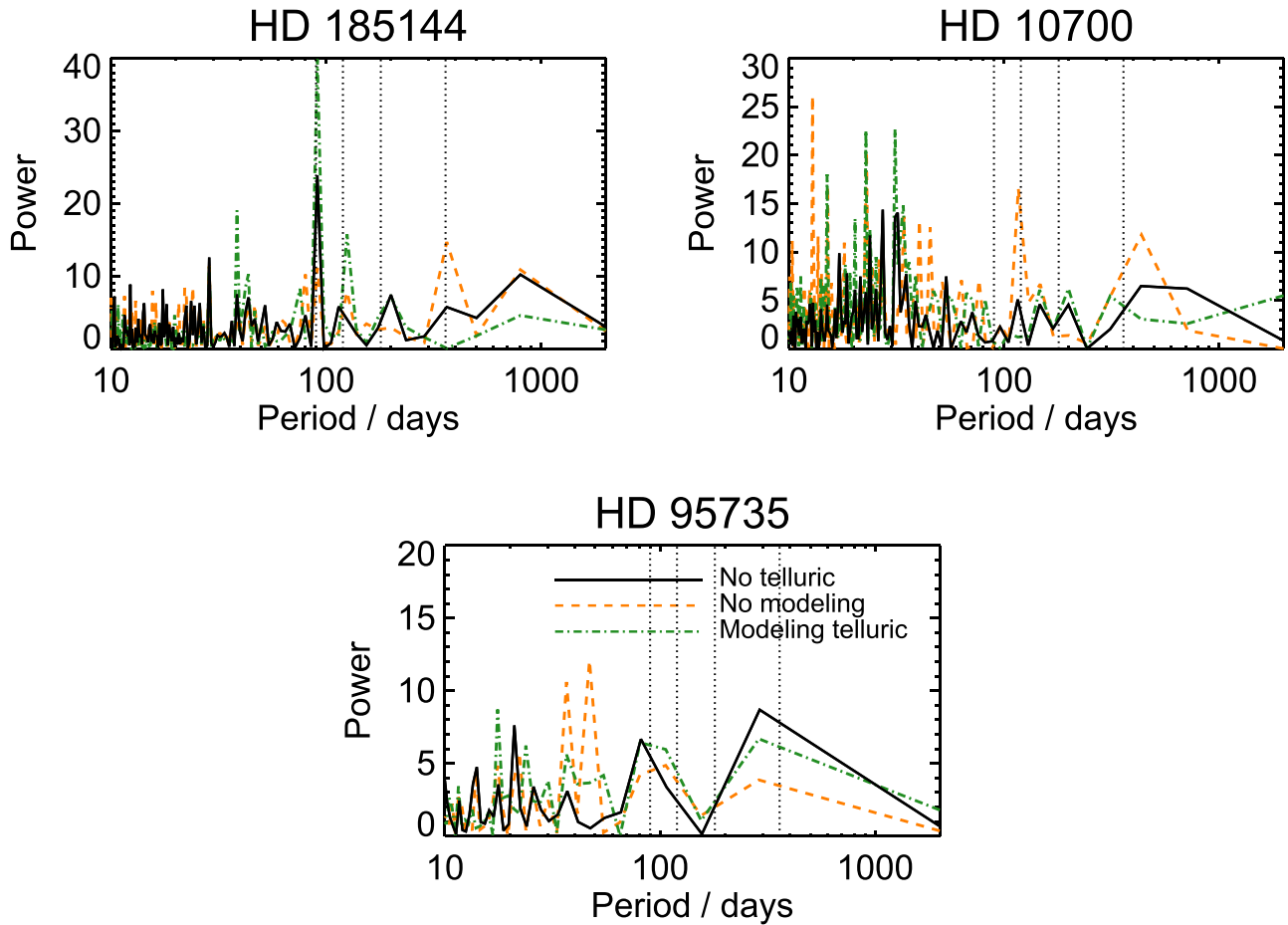
Figure 7. Periodograms of RVs from noise-free spectra at $R = 120k$ from simulated spectra without tellurics (solid black), telluric-contaminated spectra with no treatment (no modeling; dashed orange), and telluric-contaminated spectra with forward modeling of tellurics (green dotted–dashed). The vertical lines are at 90, 180, and 360 days (one year and harmonics), where systematics due to spectral contamination are most likely to occur. Telluric contamination introduces some aliasing in periodograms but not a significant amount. See Section 3.1 for more details.
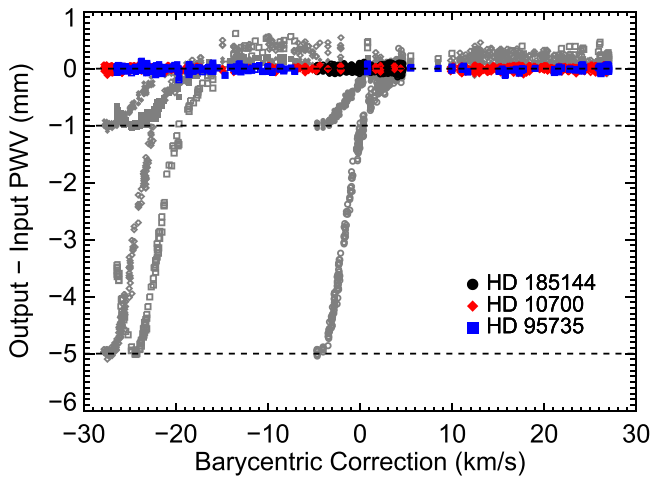
Figure 8. Difference in input and output (best-fit) PWV values vs. barycentric velocity (or barycentric correction, BC) of the star for each observation. The solid symbols are for simulations with DSST having PWV = 1 mm, and the hollow symbols are for DSST with PWV = 5 mm. All spectra have $R = 120k$ and S/N = 200. This illustrates the excellent agreement between the input and retrieved PWV values, with an rms of the residuals being 0.04 mm. The gray symbols are for simulations with DSSTs where the water lines have not been removed, and it illustrates the bias in PWV retrieval caused by tellurics in the DSSTs. The bias reaches its maximum when the BC of the star at the observation is the same as the BC of the DSST observation, i.e., when the tellurics in the observation and the DSST are at the same wavelengths. See Section 3.3 for more.

4. Fitting Tellurics in Real Observations

As demonstrated in the sections above, telluric contamination does not add a significant amount of RV errors or biases to the iodine-calibrated data, especially in consideration of the typical S/N and precision. To confirm this result, we added modeling of tellurics when extracting RVs from real Keck/HIRES data, following the procedure described in Section 2.4. The results are listed in the bottom two rows of Table 1. As expected, there was virtually no effect in terms of RV rms after adding in telluric modeling, because of the S/N and the fact that the Keck/HIRES RV systematics was probably dominated by other sources instead of tellurics (see Wang 2016 for more). Also, as expected, adding telluric modeling has no significant effects in terms of changes in the systematics in the RV–BC plane or adding spurious signals in the periodograms.

Figure 9 shows the PWV values retrieved from real Keck/HIRES observations on our three stars. The overall trends and range of values of the PWV were consistent with other site measurements for Maunakea, such as García-Lorenzo et al. (2010). To avoid biases in the retrieved PWV values caused by water lines in the DSST (Figure 8), we cleaned up the DSSTs beforehand to eliminate the water lines in them as much as possible. We visually examined the DSST spectral chunks with relatively deep water lines, and adjusted the depth and width (i.e., resolution) of the water lines to match with the observed
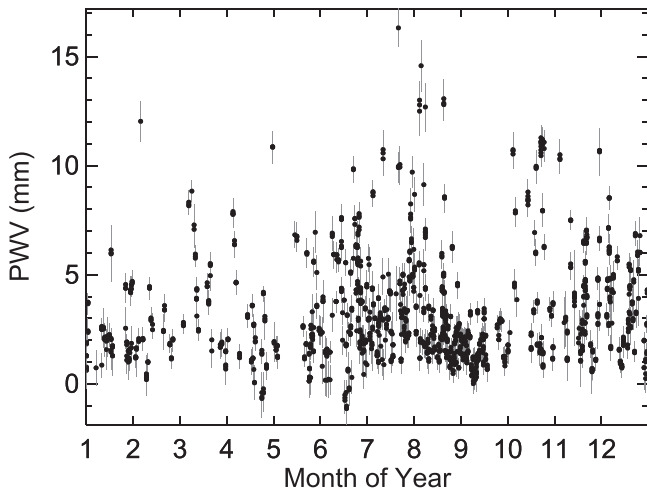
Figure 9. PWV values derived from real Keck/HIRES observations of HD 185144, HD 10700, and HD 95735 plotted against month of year. The seasonal variation and absolute PWV values are consistent with those reported in the literature (e.g., García-Lorenzo et al. 2010). The PWV values are derived from RV analyses using "cleaned DSSTs" where water lines have been divided out. The error bars are from the scatter in the PWV values derived among different spectral chunks within each observation. See Section 4 for more details.

(but deconvolved) lines in the DSST as much as possible. We then divided out the water lines given the visual best-fit width and depth (PWV) across the entire DSST to arrive a cleaned DSST. There are still some negative PWV values, which are probably a result of residuals from this division, i.e., an underestimated PWV. Unfortunately, as shown in the last row of Table 1, these cleaned DSSTs did not bring consistent improvements to the RV precision in real observations, which could be due to incomplete removal of water lines and/or additional error induced in the division process.

5. Conclusion and Future Work

Using simulated spectra similar to those of Keck/HIRES and real weather data from Kitt Peak, we have characterized the effects of telluric contamination in iodine-calibrated precise RVs. We conclude that telluric contamination introduces additional errors and systematics of the order of $10 \, \mathrm{cm \, s^{-1}}$ for iodine-calibrated RVs, similar to those previously reported for HARPS in Cunha et al. (2014) and Artigau et al. (2014). This amount of error is essentially negligible given the typical S/N and precision achieved by real on-sky iodine-calibrated RV spectra. For drier sites such as Maunakea and the Canary Islands, the effects of tellurics for iodine-calibrated RVs would be even smaller compared to the conclusion from this study based on the weather data of Kitt Peak.

We found that the adverse effect of telluric contamination is contained to $\sim 10 \, \mathrm{cm \, s^{-1}}$, mostly for two reasons: the shallow lines in the iodine region (500–620 nm) and the downweighting of the telluric-contaminated spectral region due to the low RV precision caused by tellurics.

We also added telluric lines as an additional component in the forward modeling process when extracting RVs from iodine-calibrated data, and we conclude that modeling effectively mitigates the effects of telluric contamination, both in terms of minimizing the added RV errors and in correcting the spurious periodic signals induced by tellurics. We have

demonstrated that modeling the water lines does not require prior knowledge of the PWV values, and that modeling can retrieve the PWV information somewhat accurately. However, one should caution that, when the BCs of the RV observations are close to the BC of the stellar template used in the forward modeling process, there will be significant bias in the retrieved PWV values due to telluric line blending. Unfortunately, telluric lines in the stellar template observations are very hard to remove without distorting the template and creating a net decrease in RV precision. It is possible that template-free methods for extracting RVs (e.g., Gao et al. 2016; Czekala et al. 2017) would be free of the problems caused by telluric lines in the stellar templates, which are generally derived from observations.

In practice, ignoring tellurics and not modeling them will not compromise the RV precision in iodine-calibrated RV data in any significant way, as long as all spectral regions are carefully evaluated and weighted before being combined to derive a final RV. However, for any RV program using iodine that aims at $< 1 \, \mathrm{m \, s^{-1}}$ precision, we recommend that telluric modeling be added to the forward modeling process, especially if the humidity of the telescope site is similar to or higher than at Kitt Peak (with a median PWV of ~ 5 mm). We also highly recommend that stellar template observations be taken on relatively dry nights (ideally PWV ~ 5 mm or smaller).

We notice that spectral resolution does not seem to be critical for the resilience against or the ability to recover from the impact of tellurics for iodine-calibrated RVs, which is perhaps not surprising given that both stellar lines and water lines are basically resolved with $R \sim 70$k within the iodine wavelength region. As shown in Table 1, the results for the $R = 120$k simulations do not appear to suffer less from tellurics, nor do they recover better when fitting tellurics in comparison to the $R = 70$k simulations. As the typical iodine-calibrated instruments typically only reach to up to $R \sim 130$k (e.g., the Planet Finder Spectrograph, PFS, on Magellan, Crane et al. 2010), we did not perform simulations with an even higher spectral resolution or investigate whether that would bring better performance under the influence of tellurics. However, for the next-generation spectrographs, this could be vitally important, especially for the NIR where telluric lines are copious and often heavily blended with stellar lines, especially for M dwarfs.

Importantly, though irrelevant to telluric contamination, our simulations, especially the results in Figure 3, demonstrate that the choice of resolution can be important for achieving high RV precision with iodine-calibrated instruments. In particular, for slit-fed instruments or fiber-fed instruments that are resilient against loss in seeing, choosing a higher resolution setting (e.g., $R = 120$k versus $R = 70$k) could boost the RV precision considerably given the same exposure time. This boost of RV precision has been verified in the on-going surveys using Magellan/PFS, including the RV follow-up on small transiting exoplanets discovered by *TESS* (Ricker et al. 2015). After installing a larger-format CCD with smaller pixels, Magellan/PFS switched its nominal operation mode from $R = 80$k to $R = 130$k (from using the 0."5 slit to using the 0."3 slit, or a boost in resolution and RV precision for the same S/N by a factor of 5/3). The photon loss due to a narrower slit was smaller than a factor of $(1 − 3/5) = 40\%$ given the superb seeing at Las Campanas Observatory, and as a result, the RV precision has increased considerably without requiring a

significant increase in exposure time (e.g., see the difference in old and new PFS data in Dragomir et al. 2019).[15] We thus recommend that iodine-calibrated RV programs that require high efficiency and high RV precision, such as *TESS* follow-up programs, adopt a high-resolution mode if possible. Again, for particularly challenging targets with small RV amplitudes ($\sim 1\,\mathrm{m\,s}^{-1}$), we recommend incorporation of telluric contamination to eliminate potential additional error sources and aliases in the frequency domain as much as possible.

In addition to the telluric mitigation method presented in this work (incorporating a telluric model in the forward modeling process), there are several other noteworthy methods in recent works. In particular, the data-driven method presented by Bedell et al. (2019) proves to work sufficiently well in removing the telluric contamination without requiring any prior knowledge on the telluric spectra. There are also data-driven methods for iodine-calibrated RV spectra as well, such as the work by Díaz et al. (2019) and Leet et al. (2019).

Although we argue that mitigating tellurics is not vital (though could be important) for iodine-calibrated RVs, for the next-generation RV spectrographs that aim at a precision of $10\,\mathrm{cm\,s}^{-1}$, the errors induced by telluric contamination definitely cannot be ignored. As the amplitude of the additional errors and spurious signals added by tellurics is essentially a result of the competition between the Doppler content in the stellar spectrum and the telluric spectrum, it is reasonable to speculate that future spectrographs with higher precision but extending to wider and redder wavelength ranges will face a larger amount of systematics, as the telluric absorption increases toward the red. How to mitigate telluric contamination effectively across a broad range of wavelength coverage and spectrograph setups is the topic of many on-going and future studies (e.g., Plavchan et al. 2018).[16]

ORCID iDs

Sharon Xuesong Wang (王雪淞) ⓘ https://orcid.org/0000-0002-6937-9034
Jason T. Wright ⓘ https://orcid.org/0000-0001-6160-5888
Chad Bender ⓘ https://orcid.org/0000-0003-4384-7220
Andrew W. Howard ⓘ https://orcid.org/0000-0001-8638-0320
Howard Isaacson ⓘ https://orcid.org/0000-0002-0531-1073
Mark Veyette ⓘ https://orcid.org/0000-0002-0385-2183
Philip S. Muirhead ⓘ https://orcid.org/0000-0002-0638-8822

References

Allard, F., Homeier, D., & Freytag, B. 2012a, RSPTA, 370, 2765
Allard, F., Homeier, D., Freytag, B., & Sharp, C. M. 2012b, in EAS Pub. Ser. 57, Low-Mass Stars and the Transition Stars/Brown Dwarfs, ed. C. Reylé, C. Charbonnel, & M. Schultheis, 3
Artigau, É., Astudillo-Defru, N., Delfosse, X., et al. 2014, Proc. SPIE, 9149, 5
Baker, A. D., Blake, C. H., & Sliski, D. H. 2017, PASP, 129, 085002
Baraffe, I., Homeier, D., Allard, F., & Chabrier, G. 2015, A&A, 577, A42
Baranne, A., Queloz, D., Mayor, M., et al. 1996, A&AS, 119, 373
Baron, E., & Hauschildt, P. H. 2007, A&A, 468, 255
Bean, J. L., Seifahrt, A., Hartman, H., et al. 2010, ApJ, 713, 410
Bedell, M., Hogg, D. W., Foreman-Mackey, D., Montet, B. T., & Luger, R. 2019, AJ, 158, 164
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
Butler, R. P., & Marcy, G. W. 1996, ApJL, 464, L153
Butler, R. P., Marcy, G. W., Williams, E., et al. 1996, PASP, 108, 500
Butler, R. P., Vogt, S. S., Laughlin, G., et al. 2017, AJ, 153, 208
Campbell, B., Walker, G. A. H., & Yang, S. 1988, ApJ, 331, 902
Castro-Almazán, J. A., Muñoz-Tuñón, C., García-Lorenzo, B., et al. 2016, Proc. SPIE, 9910, 99100P
Clough, S. A., Iacono, M. J., & Moncet, J.-L. 1992, JGR, 97, 15
Coelho, P. R. T. 2014, MNRAS, 440, 1027
Crane, J. D., Shectman, S. A., & Butler, R. P. 2006, Proc. SPIE, 6269, 626931
Crane, J. D., Shectman, S. A., Butler, R. P., et al. 2010, Proc. SPIE, 7735, 773553
Cunha, D., Santos, N. C., Figueira, P., et al. 2014, A&A, 568, A35
Czekala, I., Mandel, K. S., Andrews, S. M., et al. 2017, ApJ, 840, 49
Díaz, M. R., Shectman, S. A., Butler, R. P., & Jenkins, J. S. 2019, AJ, 157, 204
Dragomir, D., Teske, J., Günther, M. N., et al. 2019, ApJL, 875, L7
Feng, F., Tuomi, M., Jones, H. R. A., et al. 2017, AJ, 154, 135
Figueira, P., Kerber, F., Chacon, A., et al. 2012, MNRAS, 420, 2874
Fischer, D. A., Anglada-Escude, G., Arriagada, P., et al. 2016, PASP, 128, 066001
Gao, P., Plavchan, P., Gagné, J., et al. 2016, PASP, 128, 104501
García-Lorenzo, B., Eff-Darwich, A., Castro-Almazán, J., et al. 2010, MNRAS, 405, 2683

Han, E., Wang, S. X., Wright, J. T., et al. 2014, PASP, 126, 827
Hatzes, A. P., & Cochran, W. D. 1993, ApJ, 413, 339
Hauschildt, P. H. 1993, JQSRT, 50, 301
Hauschildt, P. H., & Baron, E. 2006, A&A, 451, 273
Howard, A. W., & Fulton, B. J. 2016, PASP, 128, 114401
Howard, A. W., Johnson, J. A., Marcy, G. W., et al. 2009, ApJ, 696, 75
Howard, A. W., Johnson, J. A., Marcy, G. W., et al. 2011, ApJ, 726, 73
Johnson, J. A., Clanton, C., Howard, A. W., et al. 2011, ApJS, 197, 26
Johnson, J. A., Marcy, G. W., Fischer, D. A., et al. 2006, ApJ, 647, 600
Latham, D. W., Stefanik, R. P., Mazeh, T., Mayor, M., & Burki, G. 1989, Natur, 339, 38
Lisogorskyi, M., Jones, H. R. A., & Feng, F. 2019, MNRAS, 485, 4804
Leet, C., Fischer, D. A., & Valenti, J. A. 2019, AJ, 157, 187
Marcy, G. W., Isaacson, H., Howard, A. W., et al. 2014, ApJS, 210, 20
Mayor, M., Pepe, F., Queloz, D., et al. 2003, Msngr, 114, 20
Mayor, M., & Queloz, D. 1995, Natur, 378, 355
Muirhead, P. S., Edelstein, J., Erskine, D. J., et al. 2011, PASP, 123, 709

Pepe, F., Lovis, C., Ségransan, D., et al. 2011, A&A, 534, A58
Pepe, F., Mayor, M., Galland, F., et al. 2002, A&A, 388, 632
Plavchan, P., Cale, B., Newman, P., et al. 2018, arXiv:1803.03960
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, JATIS, 1, 014003
Rothman, L. S., Gordon, I. E., Babikov, Y., et al. 2013, JQSRT, 130, 4
Sarazin, M. 2002, ESPAS Site Summary Series: Mauna Kea, http://www.eso.org/gen-fac/pubs/astclim/espas/espas_reports/ESPAS-MaunaKea.pdf
Sithajan, S., Ge, J., & Wang, J. 2016, AAS Meeting Abstracts, 227, 137.19
Tuomi, M., Jones, H. R. A., Jenkins, J. S., et al. 2013, A&A, 551, A79
Valenti, J. A., Butler, R. P., & Marcy, G. W. 1995, PASP, 107, 966
Valenti, J. A., & Fischer, D. A. 2005, ApJS, 159, 141
Valenti, J. A., & Piskunov, N. 1996, A&AS, 118, 595
Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, Proc. SPIE, 2198, 362
Wang, S. X. 2016, PhD thesis, The Pennsylvania State Univ.
Ware, R. H., Fulker, D. W., Stein, S. A., et al. 2000, BAMS, 81, 677
Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, A&AS, 143, 9
Wright, J. T., & Robertson, P. 2017, RNAAS, 1, 51