Draft version January 28, 2019
Preprint typeset using LaTeX style AASTeX6 v. 1.0

HIRESPRV: A COMPARISON WITH DATA PRODUCTS FROM THE CALIFORNIA PLANET SEARCH

Benjamin J. Fulton[1]

1. COMPARISON WITH CALIFORNIA PLANET SEARCH PRODUCTS

1.1. *Experimental Setup*

In order to validate the HIRESprv package we conduct a series of comparisons to the data products previously-calculated by the CPS team. The CPS spectra are processed on a Linux machine called `cadence`, while the NExScI version of the code is run on a machine called `hiresprv`. These machine names will often be used to describe the source of the data products being compared. Table 1.1 shows a summary of the computing environment on each machine. The different IDL versions on each machine was an initial concern, but we will see later that this has no impact.

Table 1. Computing Environment

Hostname	Architecture	OS	IDL Version
`cadence`	x86 64	RHEL 7.2	8.3
`hiresprv`	x86 64	CentOS 7	8.2.3

1.2. *Spectral Extraction*

We first test consistency of the spectral extraction routine which produces one-dimensional spectra extracted for each echelle order. The resulting product is a FITS image with the number of orders and the number of columns across each CCD row as the two dimensions. The flux values in this 2D image reflect the ADU flux in each order summed over several pixels centered around the peak of the spectral traces. The number of pixels in the summation varies between 6-12 pixels depending on the quality of the point spread function (PSF).

We compared a series of spectra collected on July 3 2013. The series is consecutive and consists of 3 B star observations taken with the iodine cell in, 5 spectra of the RV standard star σ Draconis (HD 185144) without the iodine cell, and 3 more B star observations through the cell. The extracted spectra data products for all 11 of these observations are exactly identical when extracted on either `cadence` or `hiresprv` (see Figure 1.

1.3. *Iodine Modeling*

The next step in the process of extracting precision RVs is to model the B star observations that bracket the iodine-free template observations. Each B star observation is modeled as the iodine spectrum transmission (as measured in the lab) convolved with the instrumental PSF. The resulting data product is called a VDIOD. Again we focus on the 6 B star observations from July 3 2013 which bracket the iodine-free template observation of σ Draconis. We compare the files produced on `comet` to those produced on `hiresprv` by comparing the FIT parameter. This parameter is the final χ^2_ν for the fit to each spectral chunk. Again, we find that the values are identical between the files produced on the two machines for all 718 spectral chunks (see Figure 2).

1.4. *Template Deconvolution*

With the shape of the PSF extracted from the bracketing B star observations discussed in §1.2 we now use that information to deconvolve the iodine-free observations of the target star. We perform a χ^2 deconvolution where a potential solution is randomly drawn, then convolved with the instrumental PSF and compared to the observation. The solution is systematically perturbed and this process is repeated until convergence is reached. The final deconvolution processes produces a "deconvolved stellar spectral template" or DSST. We find that the dconvolution process of the 5-shot iodine-free tempalte observation of σ Draconis from July 3 2013 produces identical results between `cadence` and `hiresprv`, but we note that both the template deconvolution and the iodine modeling takes $\approx 30\%$ longer on `hiresprv` despite the clock speed of the CPUs on `hiresprv` being 10% higher than `cadence`.

1.5. *Barycentric Correction*

One known area of divergence between the CPS and NExScI codes is in the calculation and correction for the motion of the observatory in the direction of the star at the photon-weighted mid exposure time of each observation. The underlying code is the same but the input coordinates of the stars are handled differently. The CPS team manually enters each target, its coordinates, parallax, and proper motion into a database that they maintain by hand. This is not feasible for the production environment on `hiresprv`, so instead we enforce a policy that ensures that the target names are

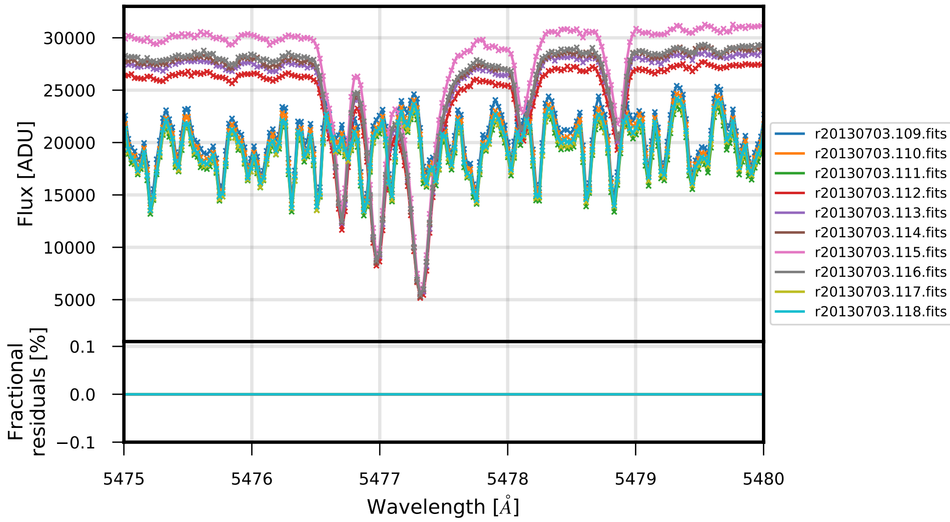

Figure 1. Comparison of the extracted spectra from `hiresprv` vs. `cadence` for a small portion of a single echelle order. The observation codes are annotated on the right. All of the plotted lines lie on top of each other and are exactly equal to zero. Spectra derived on `cadence` are represented by crosses and spectra extracted on `hiresprv` are shown by solid lines.

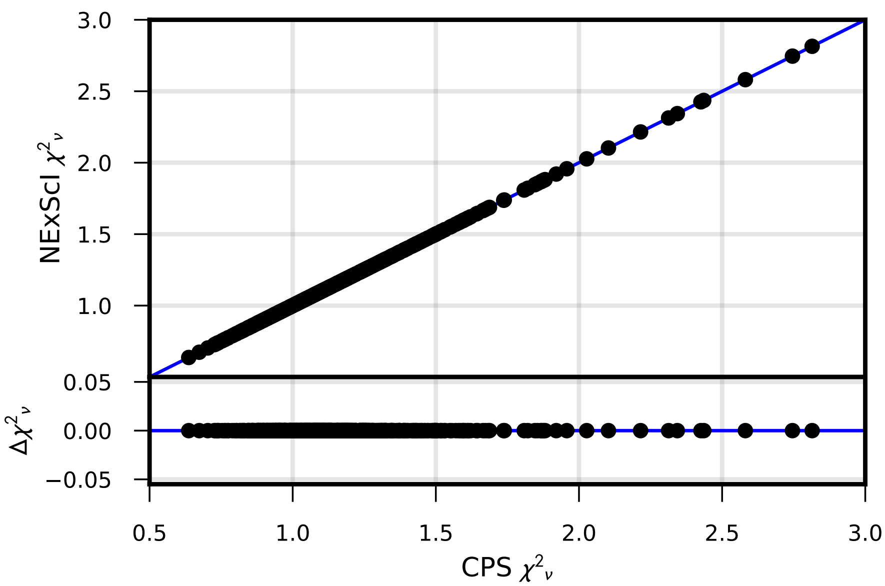

Figure 2. Comparison of a VDIOD produced with the NExScI version of the code `hiresprv` vs. the CPS version on `cadence`. Modeling of the iodine transmission functions exactly the same on `cadence` vs. `hiresprv`.

queryable on Simbad so that we can extract precise coordinates and motions from Gaia DR2. Figure 4 shows the distribution of differences between the barycentric velocities calculated on `hiresprv` vs. `cadence`. The differences are well below 25 cm/s and the vast majority of barycentric corrections (BCs) are consistent to within $\pm 5 cm/s$.

1.6. *Radial Velocity Fitting*

Since the input spectra, the PSFs extracted from the B star observations, and the DSST are all identical we can now use all of those components to fit for the radial velocity (RV) for a single observation. We forward model each observation in 718 small chunks of ≈2 Å width as

$$I_{obs}(\lambda) = k[T_{I_2}(\lambda) \cdot I_S(\lambda + \Delta\lambda)] \circledast PSF \qquad (1)$$

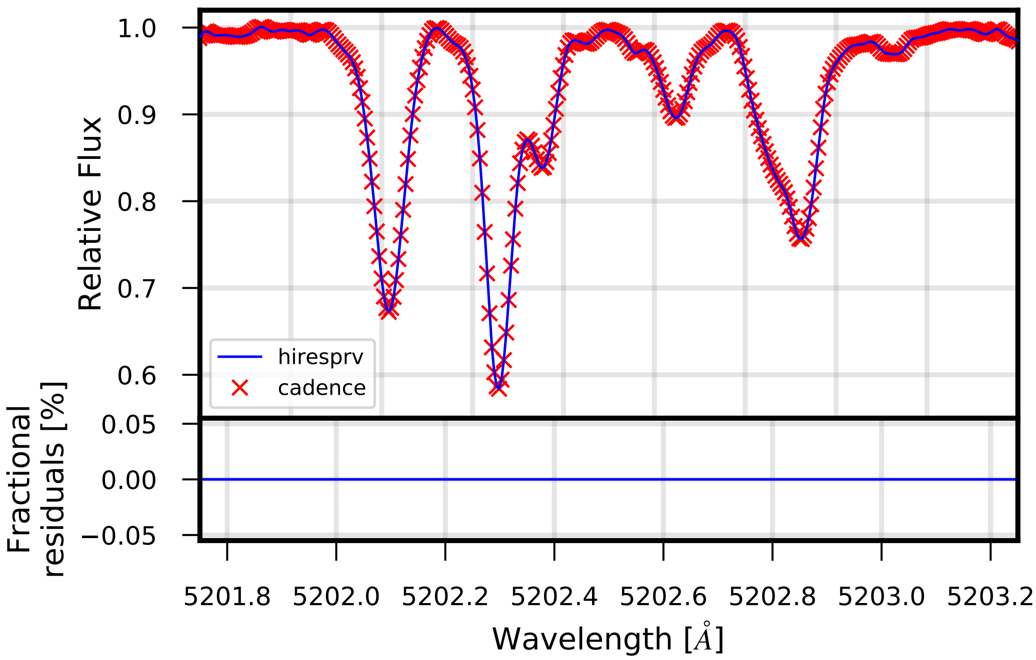

Figure 3. Comparison of a DSST produced with the NExScI version of the code `hiresprv` vs. the CPS version on `cadence`. The complex deconvolution process produces identical results.

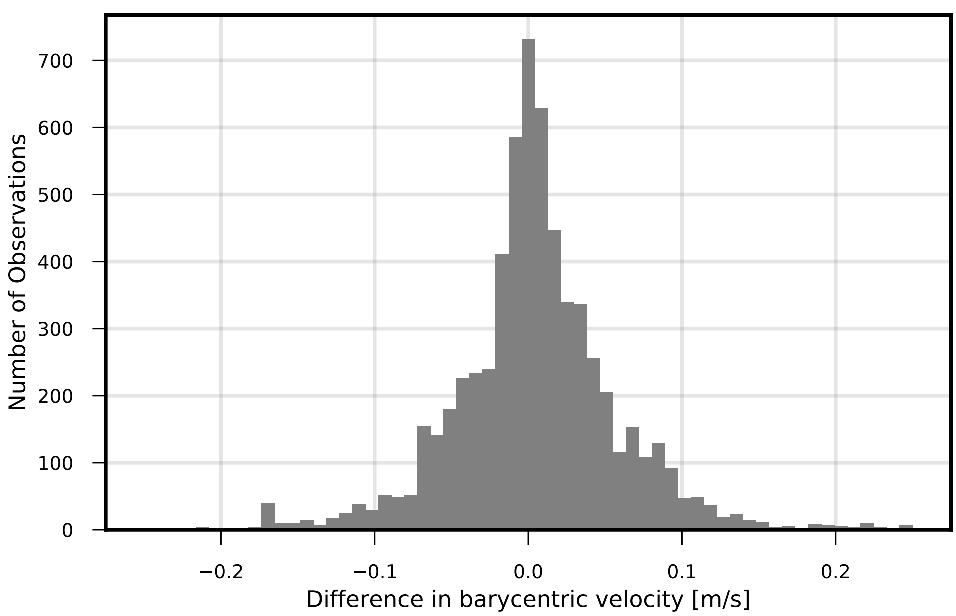

Figure 4. Histogram of the differences in the barycentric velocities calculated by the CPS code vs. the NExScI code, which is based on Gaia coordinates.

where T_{I_2} is the transmission of the iodine cell as measured in a lab, $I_S(\lambda + \Delta\lambda)$ is the DSST (perturbed by an RV that produces a wavelength shift of $\Delta\lambda$). The product of T_{I_2} and I_S is then convolved with a model of the PSF which is described by a sum of Gaussians and scaled by an arbitrary normalization factor k.

Although the difference in the BCs are extremely small relative to the typical measurement uncertainties, these tiny difference propogate into the radial velocity fits in the initial guesses given to the minimizer. Those small perturbations of the initial velocity guesses at the 5-25 cm/s level contribute to differences in the final velocities for individual chunks that can be as large as several hundred m/s. Figure 5 (left) shows that the chunk velocities as calculated on `hiresprv` vs. `cadence` are identical if the input BCs are forced to be the same. However, when the code is run using the BCs based on Gaia coordinates we do see some significant differences. Although these differences can be fairly large, it is important to note that all 718 chunks are used in the final velocity calculation as well as the statistics of each chunk as a function of time when analyzed in the context of a full RV timeseries. The maximum difference in the final mean velocity for an observation caused by a chunk that differs by 100 m/s is $100/\sqrt{718} = 3.73$ m/s and in practice these outlier chunks are heavily downweighted when constructing the final RV timeseries (Figure 6).

1.7. *Velocity Timeseries*

The final step in the creation of a radial velocity timeseries is to collect all of the chunk fits for all observations of a given star. The intra-observation and inter-observation chunk statistics are analyzed and the chunks are given weights based these statistics. Chunks that behave consistently over time and show a narrow distribution in fitted velocity values receive higher weights. The weights are used in a simple weighted mean in order to calculate a RV value for each observation.

The results of this final step are vey similar to the results we see for the individual chunk fits. Figure 7 shows that the final velocity timeseries for the 54 observations of σ Draconis analyzed here are identical if we force `cadence` to use the Gaia-derived BCs, but there are small differences in the final velocities when we compare timeseries constructed based on the Gaia BCs versus the CPS BCs. Even when using the CPS BCs the differences are ≈50 % smaller than the typical measurement uncertainties (≈0.8 m/s, or about 0.6 σ).

1.8. *Summary*

The code is functioning exactly the same on `hiresprv` and `cadence`. There are no differences in the extracted spectra, B star analysis, template deconvolution, or velocity fitting when calculated on either machine. We've adopted a new and improved method of grabbing coordinates and motions of stars based on Gaia DR2 which results in tiny differences in the BCs relative to the CPS BCs at the level of ≈5 cm/s. However, these tiny BC differences propagate into measurable differences in the final velocity timeseries at a level well below the typical measurement uncertainties.

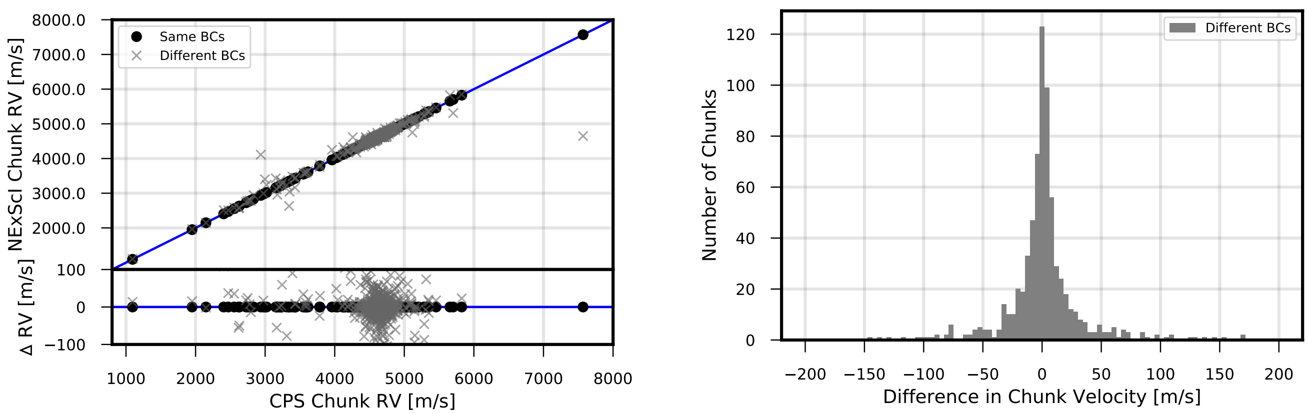

Figure 5. *Left:* Plot of `hiresprv` velocities vs. `cadence` velocities for all chunks in a single observation. The large black points are derived when forcing the `hiresprv` code to use the BCs as calculated on `cadence`. The grey crosses are derived in the default case for `hiresprv`, using the BCs based on Gaia coordinates. *Right:* Histogram of the differences between the chunk velocities calculated when using the CPS BCs or the NExScI BCs. Each of these checks contribute a maximum of $1/\sqrt{718} = 3.7\%$ toward the final velocity for the observation and outliers receive even less weight.

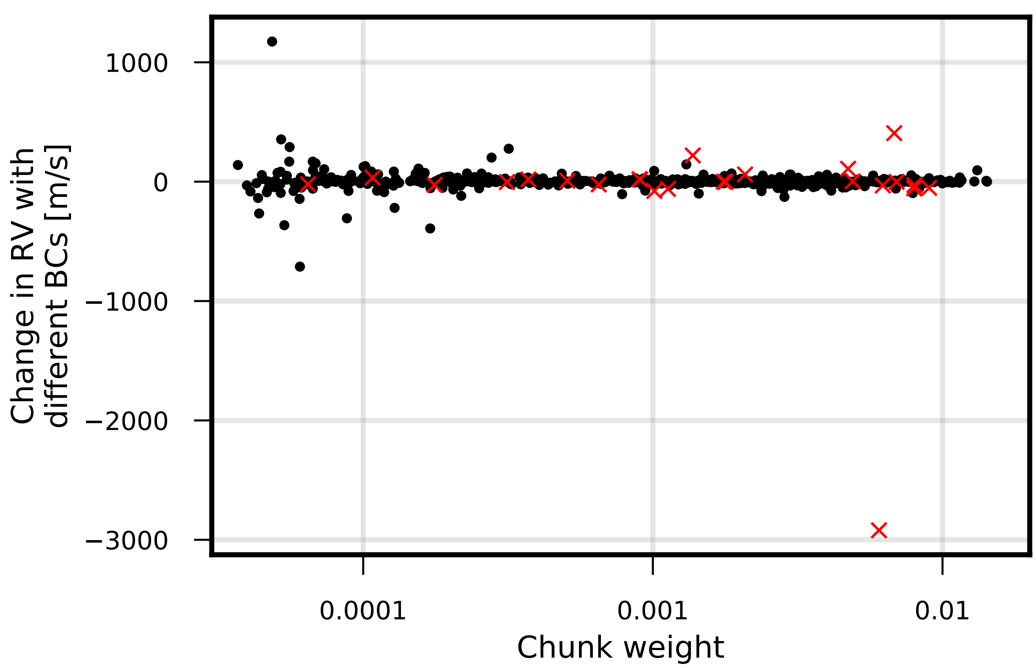

Figure 6. Difference in RVs calculated with the two different BC methods (bottom panel of Figure 5) as a function of the final weight of that chunk for the observation plotted in Figure 5. Chunks plotted as red crosses are rejected entirely. Chunks showing the largest scatter are given the lowest weight when calculating the final mean velocity for the observation.

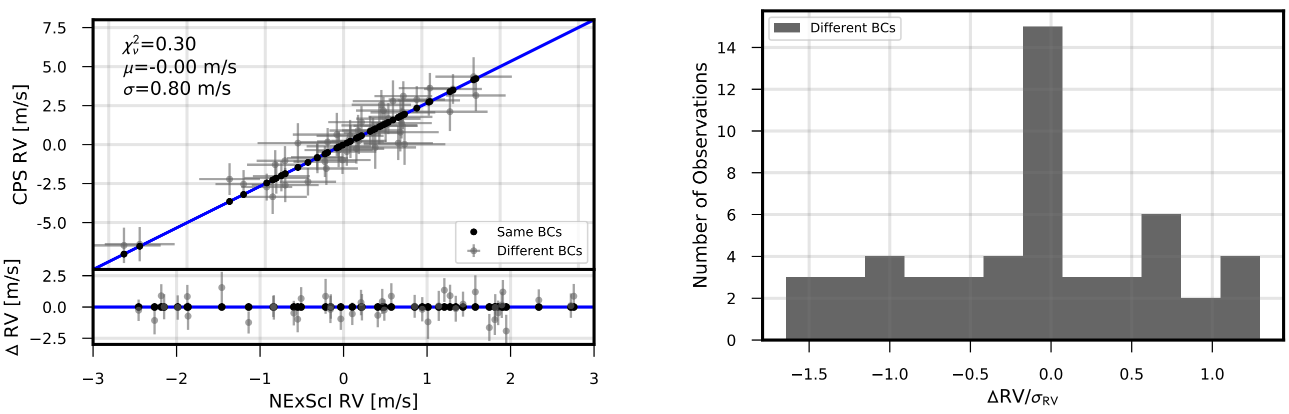

Figure 7. *Left:* Comparison of the final velocity timeseries for 54 observations of σ Draconis when computed on `hiresprv` vs. `cadence`. When `cadence` is forced to use the Gaia-derived BCs, but there are small differences in the final velocities when using the CPS BCs. *Right:* Histogram of the velocity differences divided by their measurement uncertainties.