Investigating stability of the Doppler code analysis.

Fischer - Jan 5, 2013

The goal is to check whether any variability in the unbinned tau Ceti data (**Fig 1**) comes from the Doppler analysis code (rather than instabilities in the instrument or astrophysical noise).

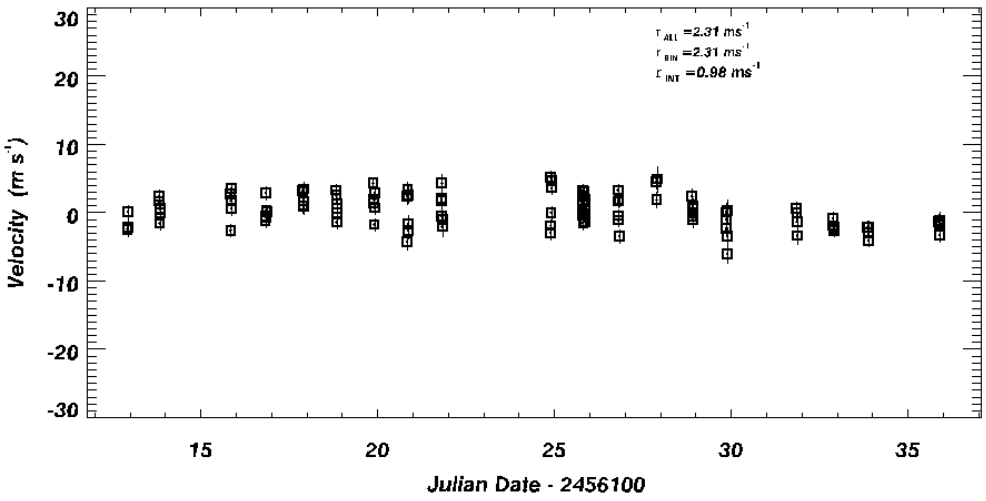


Figure 1. Unbinned tau Ceti velocities (narrow slit) in July 2012. (restore,'vstbank/vst10700.dat', x=where(cf3.jd lt 16140), velplot,cf3)

As a first step, there had been some changes to programs in the /mir7 dopcode directory so I decided to isolate the code and re-analyze the spectra. Some files had be recovered from an archived directory but the programs in /home/fischer/dop2 now seem correct. The ctio4k_init program was checked carefully and I had the following concern:

a. del_ord was set to 3
b. del_pix set to 240

This means that many chunks were included in the second-pass psf averaging (+/- 3 orders and +/- 3 80-pixel chunks) and there were signs that this was not a good thing - the SLSF peak for 160 Bstar observations was shifted from the zero point and there was variability in the modeled IP. This was changed (del_ord=1; del_pix=120) and the variability in SLSF for 160 observations decreased). Diagnostic programs are described below.

Analyzing PSF and wavelength stability.

1. ck_psf_ampl, yrmo=yrmo, bstar=bstar, star=star,hdcopy=hdcopy
- calculates a median SLSF for Bstar observations in yrmo ("1207" or Year-Month).
- extracts the SLSF pars from the yrmo bstar observations (to ampl_arr[17, n_chunks, n_vdb])
- collects wav_arr[n_chunks, n_vdb] and disp_arr[n_chunks, n_vdb] and jd_arr[n_vdb]
- calculates del_w and del_d (shifts in time of chunk wavelengths and dispersion)

ck_psf_ampl will return vdb a strarr of names in yrmo, then pst='cdadgb'+strmid(vdb, 5,strlen(vdb)-5) can be input to dop_diagnostics, cdbfile=pst[0],psfstack=pst,/ctio to generate a 3x3 multiplot sampling the psf over the 2-D chunk format.

2. dop_diagnostics, cdbfile=cdbfile, wav_cont=wav_cont, disp_cont=disp_cont, rms_map=rms_map, psfstack=psfstack, hdcopy=hdcopy, ctio=ctio
- reads a cdbfile and checks wavelength and dispersion continuity across the chunks
- checks goodness of model with a surface map of rms fits for the 2-D chunks
- optionally takes a stack of cdb files (psfstack) and checks variability in the PSF models.

These programs calculate the median SLSF and the standard deviation at each oversampled point. Nine chunks in the psf_arr are plotted in **Figure 2**, showing individual SLSF models for 160observations and a median with stddev (overplotted in blue).

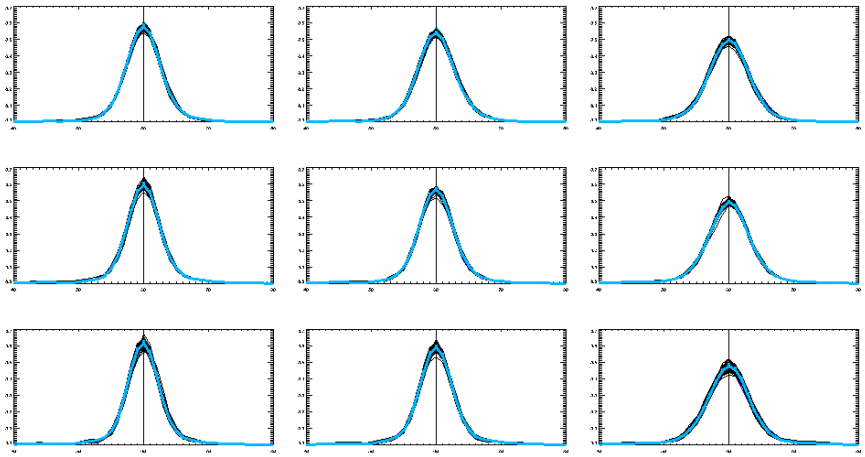


Figure 2. SLSF models for 9 chunks distributed over the 2-d format for Bstars observed with the narrow slit in July 2012. The blue line is a median PSF and the standard deviation at each oversampled pixel are indicated as error bars. CHIRON with the octagonal fiber seems to be very stable. (ck_psf_ampl, yrmo='1207',/bstar then pst='cdadgb'+strmid(vdb,5,strlen(vdb)-5) then dop_diagnostics,cdbfile=pst[0], psfstack=pst, /ctio)

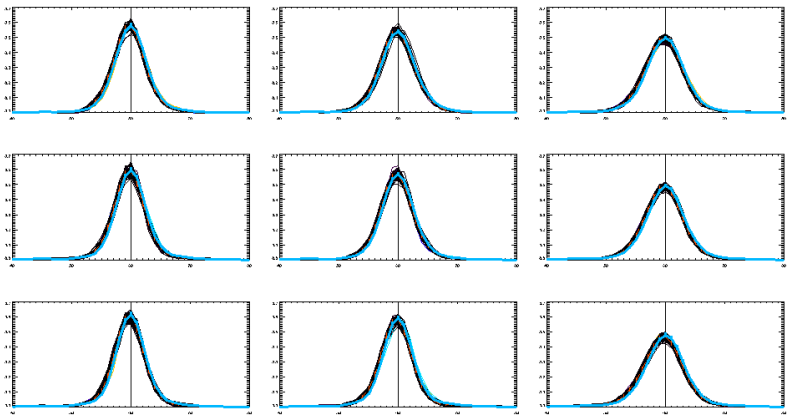
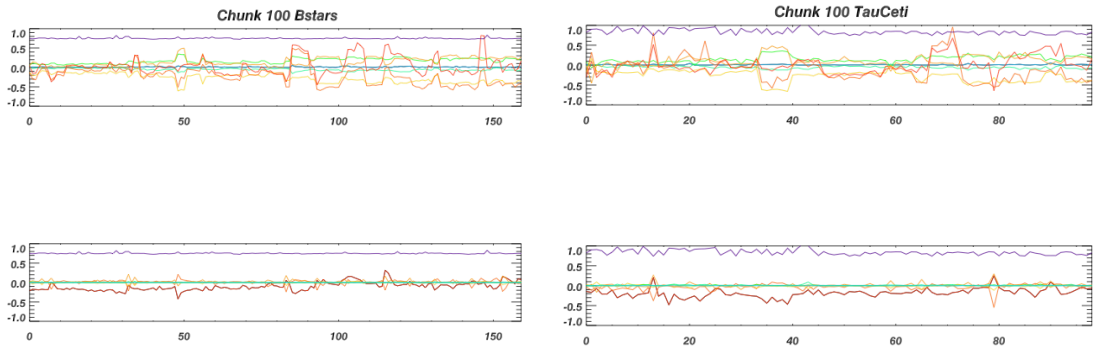


Figure 3. SLSF models for 9 chunks distributed over the 2-d format for 109 observations of tau Ceti, all taken with the narrow slit in July 2012. The blue line is the median and stddev PSF for the Bstars (Fig 2). In comparison to the 160 models for the Bstars, tau Ceti is more variable and the centroid is displaced. (ck_psf_ampl, yrmo='1207',star='10700' then pst='cdadgb'+strmid(vdb,5,strlen(vdb)-5) then dop_diagnostics,cdbfile=pst[0], psfstack=pst, /ctio)

- The SLSF for Tau Ceti (**Fig 3**) is displaced systematically relative to the SLSF for the Bstar. Here is my working hypothesis: In the zeroth pass of the Doppler code, I use a Gaussian for the PSF just b/c I don't want cross talk between PSF pars and the wavelength zero point or Doppler shift. That tweaks the wavelength zero point for the star based on a centered Gaussian PSF and would result in a systematic shift in wavelength. Then, when I fit for the PSF in the star, the centroid of the asymmetric PSF is displaced by a quarter of a pixel when the wavelength is refit in the first pass. Possible support for this hypothesis: a zeroth pass offset in the wavelength that is corrected in the first and second passes. The fix: still use the ipcf.dat system for wavelength soon but in the zeroth pass use the monthly median PSF for centering and as an initial guess. This could significantly improve the stability of the model.

- the SLSF for Tau Ceti (**Fig 4**, right) shows more variability than the SLSF for the Bstars (**Fig 4**, left). All observations were taken over the month of July, 2012). Fig 4 was generated with psfb.pro and psf2.pro).



- *Figure 4. Difference in 17 PSF pars from the first one (in time) for a series of 160 Bstar observations (left) and 109 Tau ceti observations (right) obtained in July 2012 to see variability with time. The plots are for a random chunk [100] to the left (top sets) and right (bottom sets) of the central Gaussian. (ck_psf_ampl, tag='adg',yrmo='1207',/bstar or star='10700' - writes ampl_arr.dat; then psf2)*

- All observations are taken with the narrow slit, so there is no reason for the Tau Ceti observations to have a larger FWHM (1.0 instead of 0.8) for the central Gaussian. This might be compensated by larger negative amplitude Gaussians near the center for the SLSF models of the star.

In addition to understanding the SLSF, we would also like to understand the stability of the wavelength and dispersion modeled in our data. Again, the Bstars provide the ideal objects for this check because there is no Doppler shift and the SNR of the observations is high. In ck_psf_ampl.pro, the wavelengths and dispersion for each chunk are extracted for the yrmo set of observations. In **Figure 5**, delta_wavelength for chunk[0] is shown as a function of the JD of the Bstar observation. There are multiple dots at each JD because multiple Bstar observations were made and shows a slow drift over time (should be correlated with temperture inside of CHIRON). **Figure 6** shows the same calculation, but for dispersion.

The vacuum enclosure for the echelle grating (CHIRON upgrade) was supposed to stabilize the pressure and hence the dispersion and it appears to have achieved this goal. Although the dispersion changes little, the spectrum is still drifting on the detector (**Fig 5**). The spread in

delta_wav shown (y axis, Fig 5) is expected since it reflects the difference for 760 chunks at one point in time. It does not make sense to adopt a wavelength soln from the median yrmo bstar since (**Fig 5**) the wavelength on each pixel is changing, however the dispersion could be set with a parinfo limit of +/- 0.0002 from the median yrmo chunk dispersion.

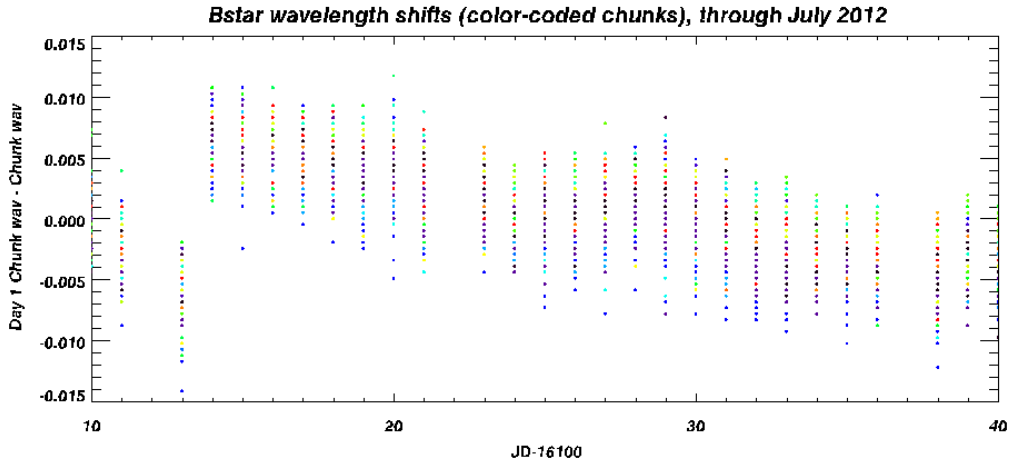


Figure 5. For Bstar observations taken in July 2012, the first wavelength solution of chunk[0] drifts with time. (ck_psf_ampl, yrmo='1207',/bstar generates this figure)

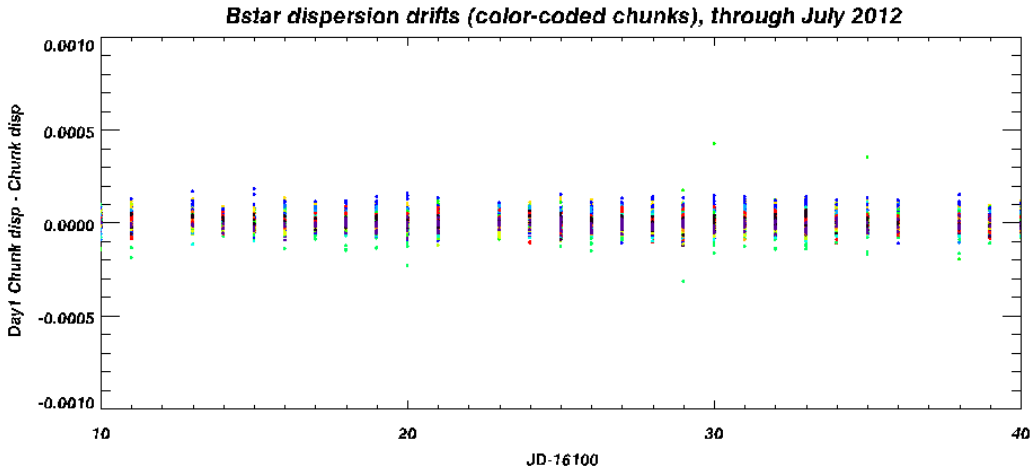


Figure 6. For Bstar observations taken in July 2012, the dispersion is stable with time to a level of 10e-4. (ck_psf_ampl, yrmo='1207',/bstar generates this figure).

The median dispersion for all 160 Bstars is shown for each chunk in **Figure 7** (left) with the nearly invisible oploterr of the stddev of the dispersion values. Zooming in, the standard deviation of the dispersion for all of these observations is shown in **Figure 7** (right). There is certainly more scatter in the dispersion solution for the Bstars in chunks 600 - 760, which corresponds to the last four orders, with wavelengths form 5865 to 6125 Angstroms. The SNR has not changed in these four orders relative to bluer orders. Two physical differences are that

the depth of the iodine lines is gradually attenuating and telluric lines are beginning to appear. **Figure 8** plots the Bstar-iodine spectrum for Order 28 (black) and the next 4 orders are overplotted in blue, green, yellow and red and show the decrease in the I2 line depth.

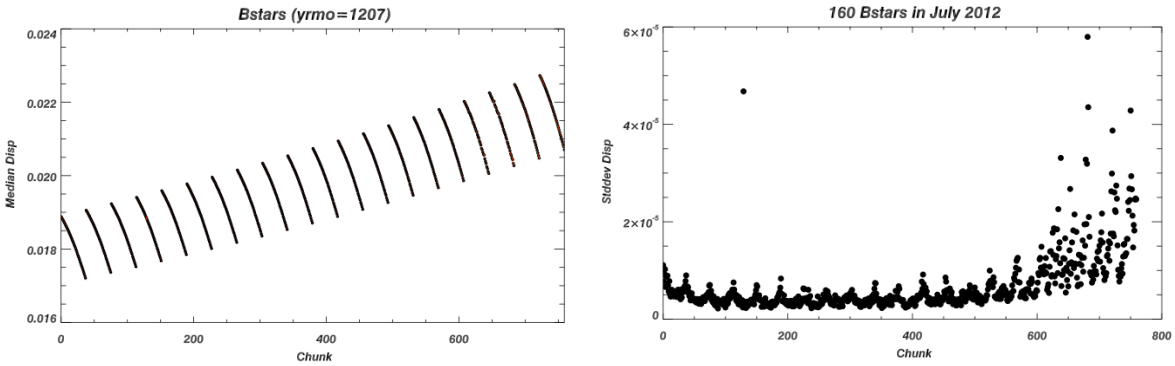


Figure 7. (left) Median dispersion in a set of 160 yrmo=1207 Bstars is shown for each chunk with the stddev overplotted as (nearly invisible) errors. (right) Zoom in on the stddev for the dispersion in the 160 Bstars. Chunk numbers larger that 600 correspond to spectra in orders 29, 30, 31, 32 and wavelengths from 5865 - 6125 Angstroms. (dop_bstar_psf, yrmo='1207',tag='adg' then plot, meddisp,ps=8 and plot, disp_rms,ps=8)

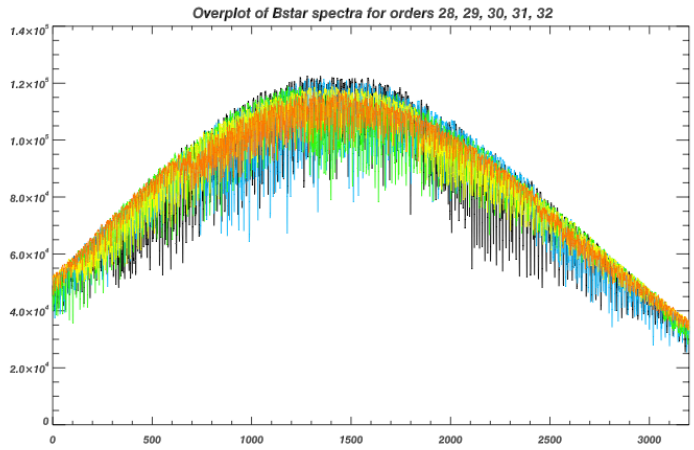


Figure 8. For one Bstar (HR8425_achi120710.1137) orders 28 - 32 are plotted in black, blue, green, yellow and red respectively. The SNR is nearly constant, so the likely reason for the increase in scatter (Fig 7, right) in the dispersion soln is that the iodine line depths are beginning to attenuate in the last 4 orders. In addition, tellurinc lines begin to appear in these last 4 orders. (rdsk,st,'/tous/mir7/iodspec/achi120710.1137,1 and plot, st[,28] with overplots of orders 29-32)*

Conclusions from diagnostic checks of SLSF, wavelength and dispersion: The SLSF models and the dispersion can (should) be constrained by the Bstar observations for fiber observations. At the very least, the median value in a given month (yrmo) should be used as a singular initial guess for the Doppler code when analyzing all yrmo program stars. The parinfo file should set limits for the dispersion that reflect the stability shown in **Fig 7,** as well as an appropriately small step size.

Revisions to the Doppler code

Taking into account the analysis above, dop_bstar_psf, yrmo=yrmo, tag=tag, plot=plot, hdcopy=hdcopy now calculates the median SLSF for each chunk in a yrmo set of Bstar observations. The median SLSF is fit with a sum of Gaussians and the amplitudes are used as inital values. **Figure 9** shows the median SLSF (black) and the model fit (red dashed) with each of the individual Gaussian components (and the central position, or psfpix, for the flanking Gaussians). The psfpix range appears wide enough to model the wings of the SLSF.

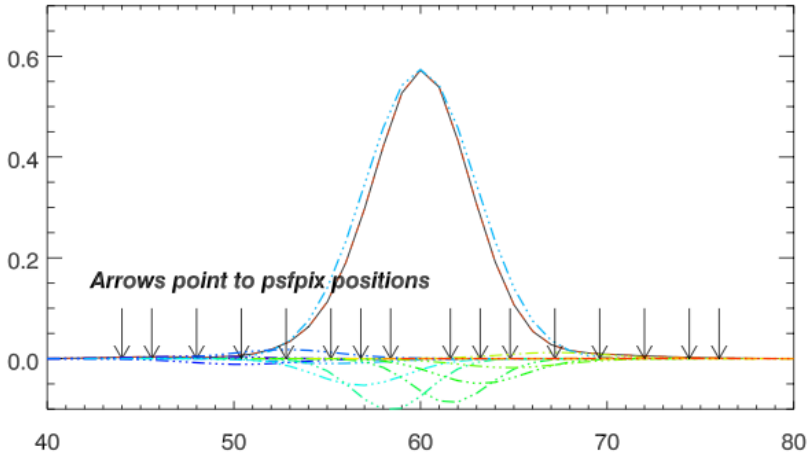


Figure 9. The median SLSF for one chunk and a set of 160 yrmo Bstar observations (black line) is overplotted with the (red-dashed) model derived from the sum of Gaussians. The central Gaussian in the SLSF model is shown with a blue dot-dashed line. The amplitudes of the flanking Gaussians (dot-dash colored lines) comprise 16 free parameters. The center positions of the side Gaussians are indicated with arrows.

The dop_bstar_psf program also calculates the median dispersion and the stddev of the dispersion. The output of this program is a cdbfile (cd+tag+'bAVG_'+yrmo). If the keyword yrmo is supplied in dr_obs.pro, then the neighboring (in time) bstar chunk_arr is restored (for the wavelength soln), however, the ip is overwritten with the median SLSF with the best fit Gaussian amplitudes. The following programs were edited to add the (optional) yrmo keyword: dr_obs, dr_star, dop_main, dop_chunk_setup.

tag: **adg** (original) is a careful analysis with reduced number of chunks for psf averaging (**Fig 1**) currently in /tous/mir7/files_df/archive_adg/

tag: **adg** uses the median SLSF as a starting point (no change to wavelength or dispersion and still fit a Gaussian in the pre-pass. This modifies the width of the central Gaussian in the first pass. The velocities for July show slight reduction in the RMS (2.08 ms-1 compared to 2.27 m/s)

tag: **adg** now uses a median SLSF as a starting point and this is fixed in the pre-pass. (all old adg files on my home iMac have been deleted). Several modifications to programs: dop_marq, dop_parinfo, dop_fit so that the central Gaussian in the first pass is not modified (derived from fitting the median yrmo SLSF). Remarkably consistent and lower chisq fits for the first pass were seen with the adh analysis and the RMS dropped.

The number of “bad” nights with high RMS scatter is another diagnostic of the robustness of Doppler models. The stddev of the individual nightly velocities (cf3.mnvel) are calculated in scatter_nightly.pro. The adh analysis provided the best results with consistently lower RMS scatter in a given night.

A key question is this: can we find any correlations (PSF pars, wavelength discontinuties, dispersion errors) that correlate with the velocities that are more than 2 sigma from the mean nightly velocity?

It is difficult to assess whether RMS scatter arises from instrumental instability, astrophysical noise, or low mass planets. Tau Ceti has a two-week window of observations with remarkably low RMS. While this interval is arbitrary, it reflects instrumental stability over a moderate time interval that has not been demonstrated with the Lick/Hamilton or Keck/HIRES spectrometers. The Doppler results are summarized in **Table 1** and plots of the Doppler velocities in a 2 week window are shown in **Figure 10** (original and revised adg analysis) and **Figure 11** (adh analysis, including the binned nightly velocities).

Table 1. Summary of Doppler analysis for 10700

tag	single measurement precision [m/s]	2-week RMS [m/s]	1-month RMS [m/s]	Number of nights with RMS < 2 m/s, > 2 m/s, > 3 m/s
adg (orig)	0.94	2.07	2.27	13, 5, 2
adg (init G.)	0.93	1.82	2.08	9, 4, 2
adh (fix SLSF)	0.90	1.59	2.11*	17, 2, 1
new adg				

**monthly RMS drops to 2.01 if the one night with greater than 3 m/s RMS is dropped.*

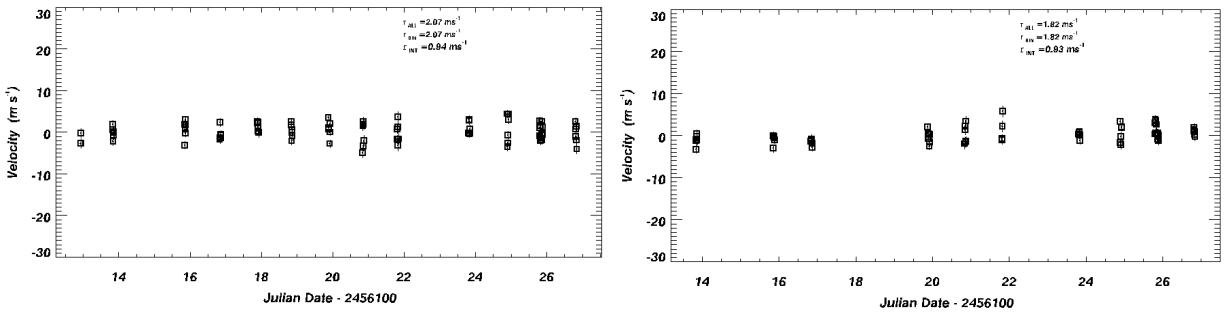


Figure 10. (left) the original adg analysis (full month shown in Fig 1) for 2 weeks of Tau Ceti with individual errors of 0.94 m/s and RMS of 2.07 m/s. (right) the revised adg analysis for the same data set (but with central Gaussian in pass-0 from the fit to the median PSF) shows nightly error of 0.93 m/s and improved RMS of 1.82 m/s .

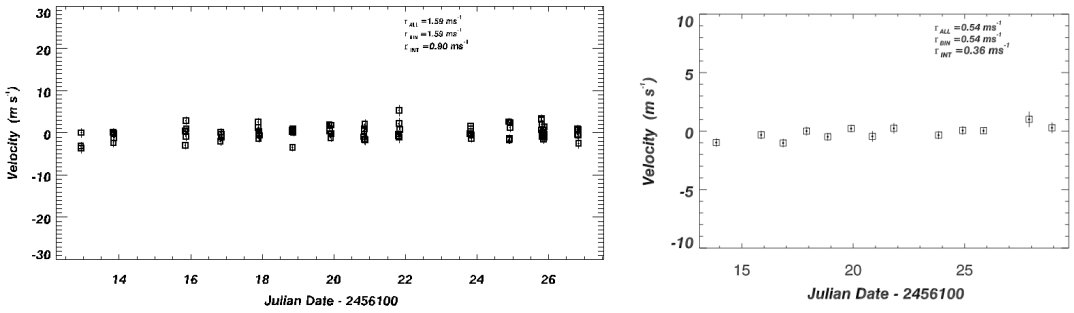


Figure 11. (left) adh analysis, fixing the SLSF to the fitted value for the monthly median Bstar SLSF shows slight improvement in the RMS of 1.59 m/s and internal errors for the unbinned data. (right) binned adh data show nightly precision of 36 cm/s and RMS of 54 cm/s.

Conclusions from Doppler analysis:
Because it is hard to know whether RMS scatter arises from instrumental stability or from the star (astrophysical sources or dynamical velocity shifts from planets) we focused on a 2-week stretch of Tau Ceti data from July 2013 that was particularly stable, but then considered the full month of analyzed data.

There is significant improvement in the *first pass* chisq fits with the adh analysis. The first pass fit is not necessarily a critical factor - we see that the chisq fits recover in the second pass of the adg analysis. However, it implies that we are close to the best fit with our initial guess parameters. We find that the overall goodness (chisq fits, individual errors, nightly and long term RMS scatter) are best for the adh analysis. Although the RMS of the adh data for the entire month of July 2012 was slightly worse than for the adg analysis (2.11 and 2.08 m/s respectively), if the one night where the RMS was greater than 3.0 m/s is dropped, the RMS over one month drops to 1.97 m/s.

To do:
1. try to find a correlation in model parameters with velocities that have > 2-sigma scatter on a given night.
2. DONE: Create dsst's using the median monthly (yrmo) Bstar SLSF instead of the bookend iodines. The Doppler analysis using this dsst gave worse results either b/c the dsst used an incorrect SLSF or because smoothing was implemented for the average wavelength and dispersion and that was not being matched.
3. DONE: run all Bstars for yrmo solutions and check the stability (SLSF, wavelength, dispersion) over longer times (all narrow slit observations done with adg)
4. DONE: run the data analysis for Tau Ceti observations made with the regular slit (need a new Bstar PSF description). Running arg on yale machines (change in number of free pars for pass0). Flaw in the analysis: Bstar observations with regular slits were not obtained.
5. run all of the CHIRON Doppler analysis.

Comparison of narrow slit and regular slit observations

The spectral resolution is measured from the thorium-argon lines; the median width for Gaussian fits to many lines is adopted as the FWHM, then R = (lambda/dlambda) / FWHM. The resolution for the narrow slit has R=136,000 and for the regular slit, R=95,000. The resolution of the regular slit is therefore quite high (compared to Lick/Hamilton or Keck/HIRES) and the exposures are either shorter by about 70% or yield higher SNR.

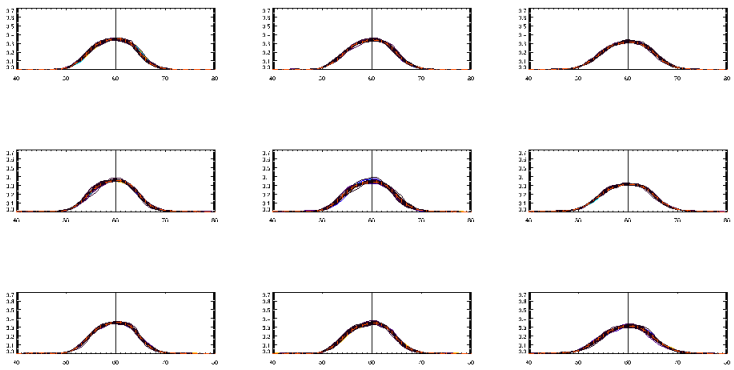


Fig 12. Bstar SLSF dop_bstar_psf, yrmo=yrmo, tag=tag, plot=plot, hdcopy=hdcopy for the regular slit are much broader by eye.

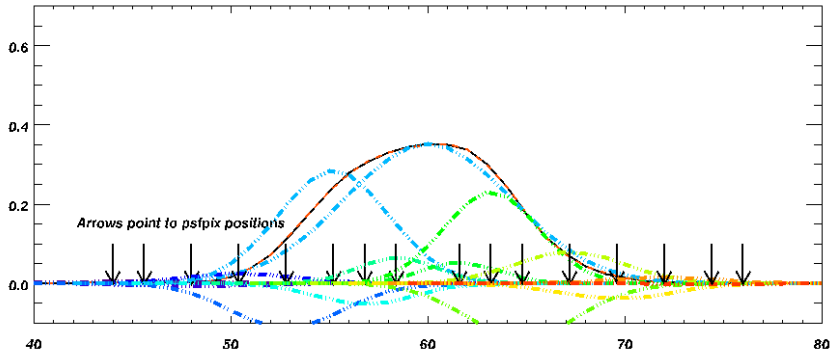


Fig 13. After the median yrmo SLSF is determined for each chunk dop_bstar_psf, yrmo=yrmo, tag=tag, plot=plot, hdcopy=hdcopy.

New program, ck_wdif, star=star, tag=tag, yrmo=yrmo, all=all finds the change in the model wavelength and dispersion from the intial guess. This highlighted some odd behavior - the Bstar looked reasonable (**Fig 14**), however there were large systematic shifts for 10700 (**Fig 15**).

Jan 13, 2013 Change: in dop_chunk_setup, change the initial wavelength and dispersion from the dsst soln to the Bstar soln. This *should* be OK, since the starting pixel is shifted by init_z and the w0 and disp are recalculated and may take care of the shift in wavelength... or not. running 128621; need to re-run 10700 (consistency, maybe also the arg runs) and 128620.

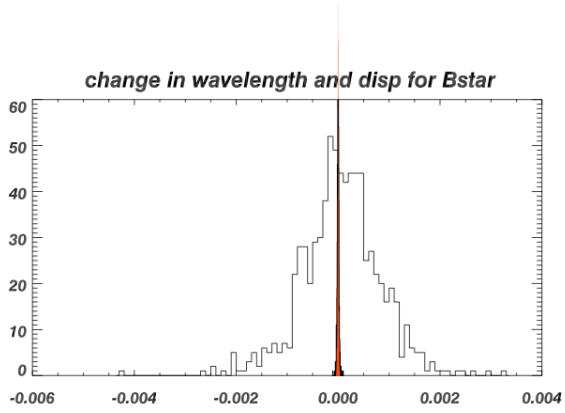


Fig 14. Difference in the final model wavelength and the initial guess (filled in dop_chunk_setup) and (in red) the difference for the intial to final model for dispersion (ck_wdif,star='HR',tag='adg',yrmo='1207')*

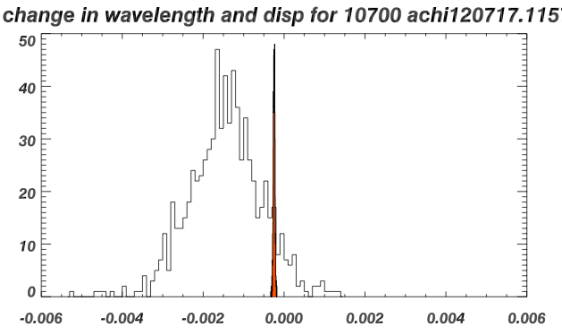


Figure 15. Same as Fig 14, but for 10700. (ck_wdif,star='10700',tag='adg',date='120717')

dsst	comments
dsst10700adg_achi120710.1141.dat	dop_dsst, bookend l2 for psf and wavelength
dsst128621adg_achi120606.1135.dat	ditto
dsst128620adg_achi120606.1124.dat	ditto