Analysis of Per Fiber Orderlet-Level Radial Velocity Using the KPF Pipeline

Dhriti Vasu

May 1 2025

Overview

This project analyzes per-order radial velocity (RV) data from the Keck Planet Finder (KPF) using etalon. A new method was implemented to extract and plot RVs order-by-order for each science fiber (SCI1, SCI2, SCI3) across both green and red detector chips. The goal was to assess RV stability and detect systematic trends or fiber-level inconsistencies using data in October 2024.

The KPF timeseries database was queried using a custom plotting function, and all results were generated using a Jupyter Notebook. The notebook is divided into six sections—one for each fiber/chip combination—followed by a final table of results.

Method

A method named `plot_rv_per_fiber_wavelength` was implemented in `analyze_time_series.py` to query the KPF timeseries database and generate offset plots of median-subtracted RV values. The method loops over the relevant RV columns for a given fiber and chip, filters by CCF weight, and creates scatter plots vertically offset for each orderlet so that the user could view all orderlets on one plot. These plots help reveal changes in RV over time, noise in specific orders, and any long-term shifts.

Per-Fiber Analysis and Plots

Etalon RV for SCI1 Fiber, Green Chip

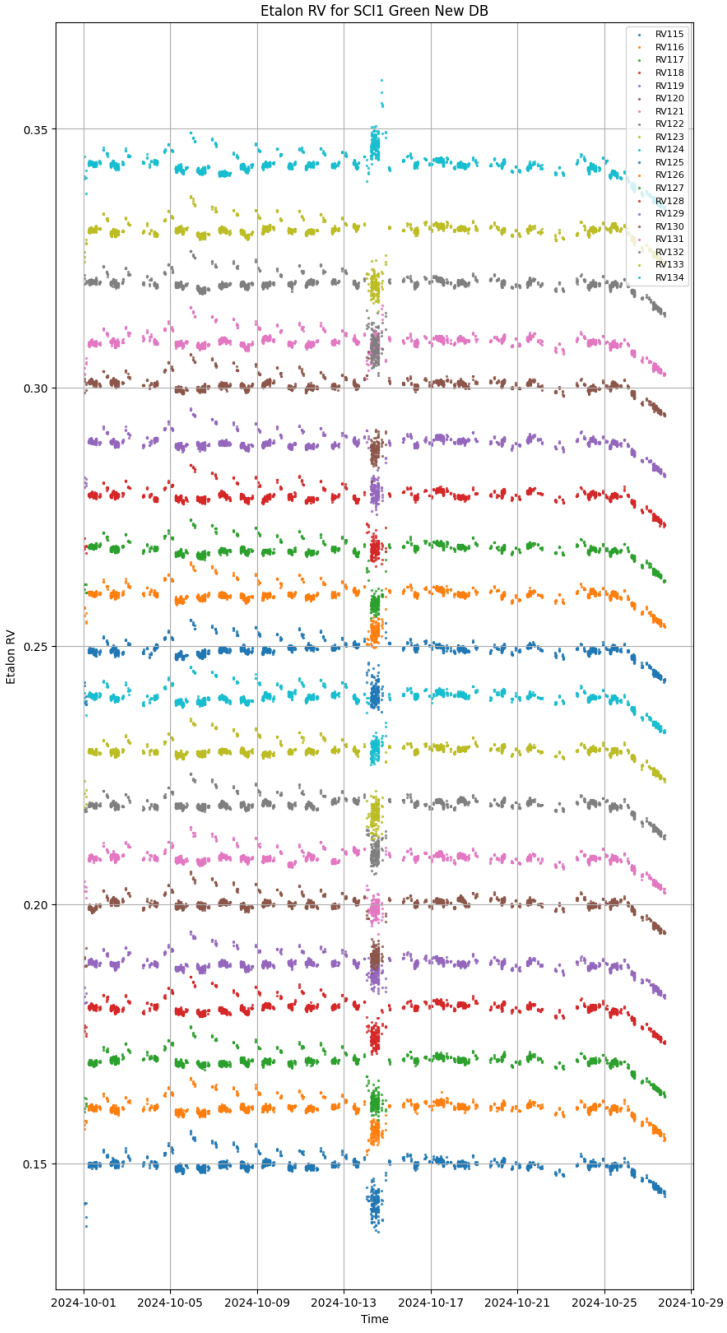

Figure 1: Order-by-order RV time series for SCI1 on the green chip (orders 100–134). Most orders show consistent, flat behavior around a median value with small scatter.

Etalon RV for SCI1 Fiber, Red Chip

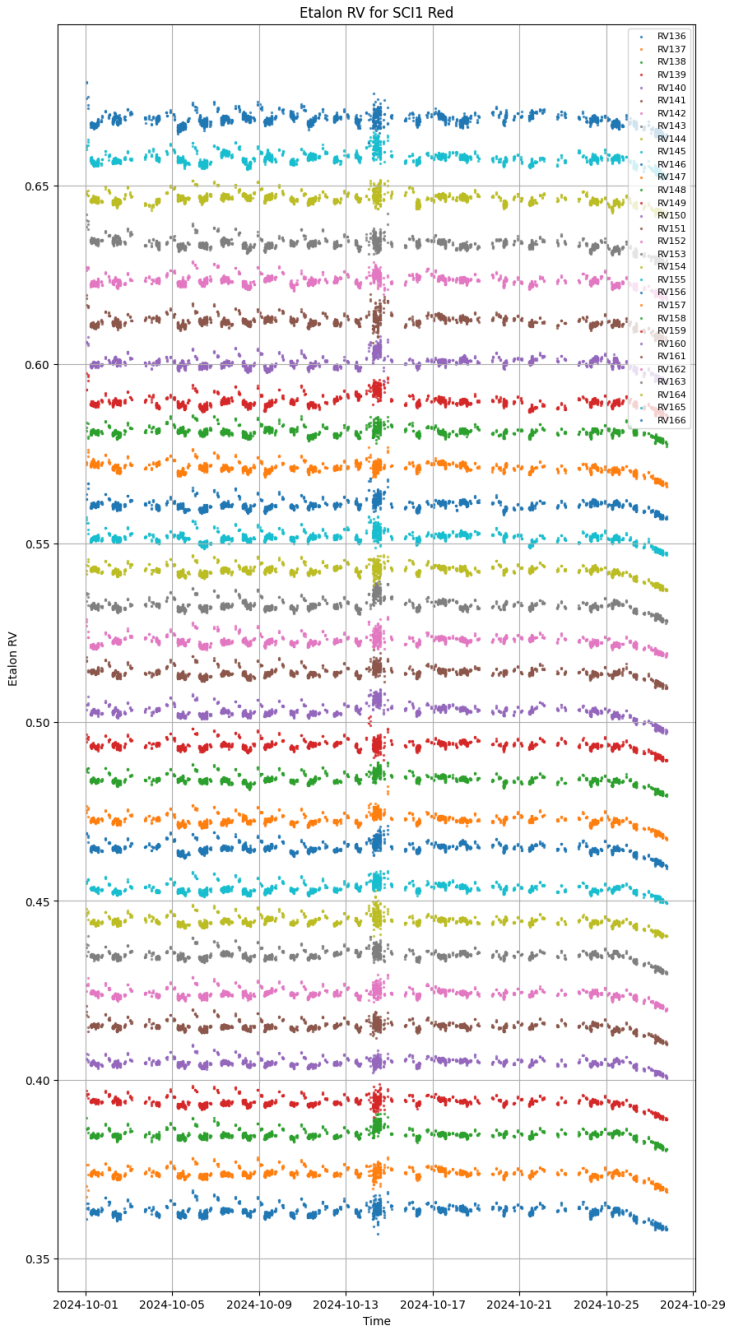

Figure 2: Red chip orders for SCI1 (orders 135–166). Most orders show consistent, flat behavior around a median value with small scatter.

Etalon RV for SCI2 Fiber, Green Chip

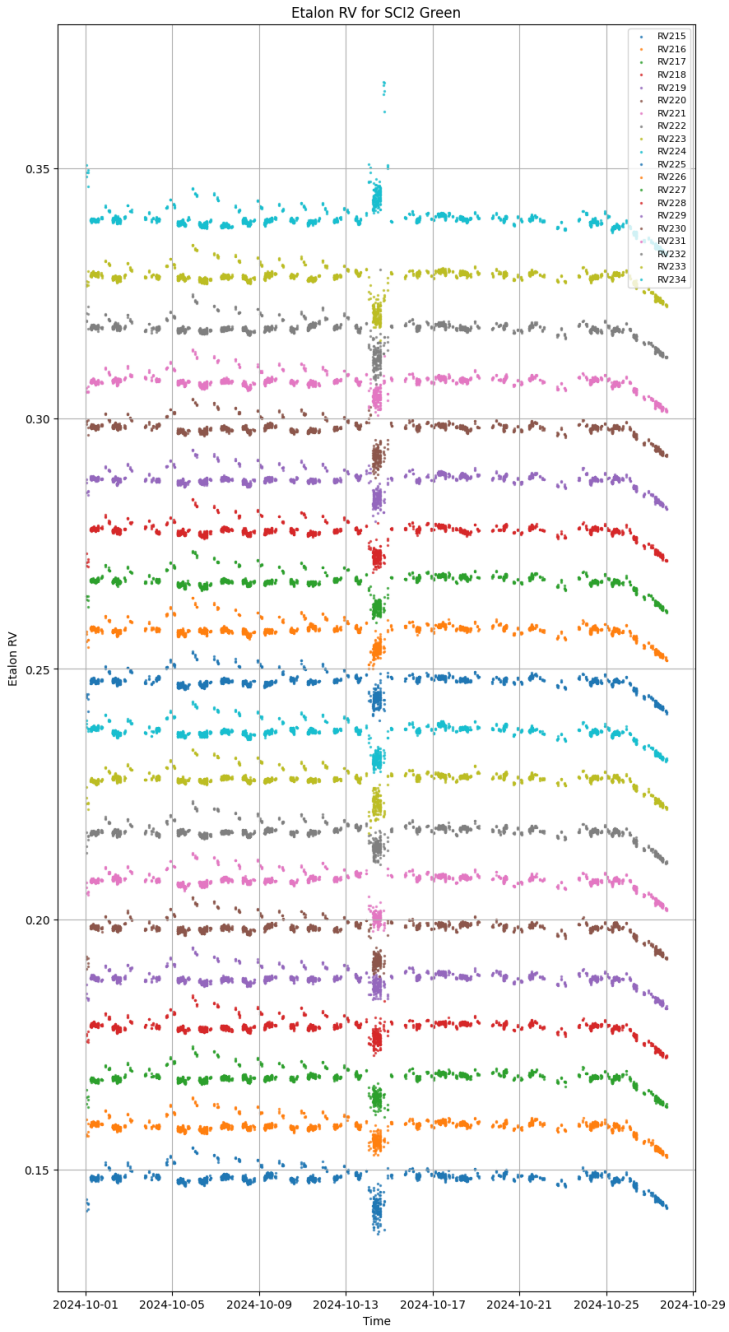

Figure 3: SCI2 green chip (orders 200–234) shows tight clustering of most orderlets, though with slightly more scatter especially around October 15th compared to SCI1.

Etalon RV for SCI2 Fiber, Red Chip

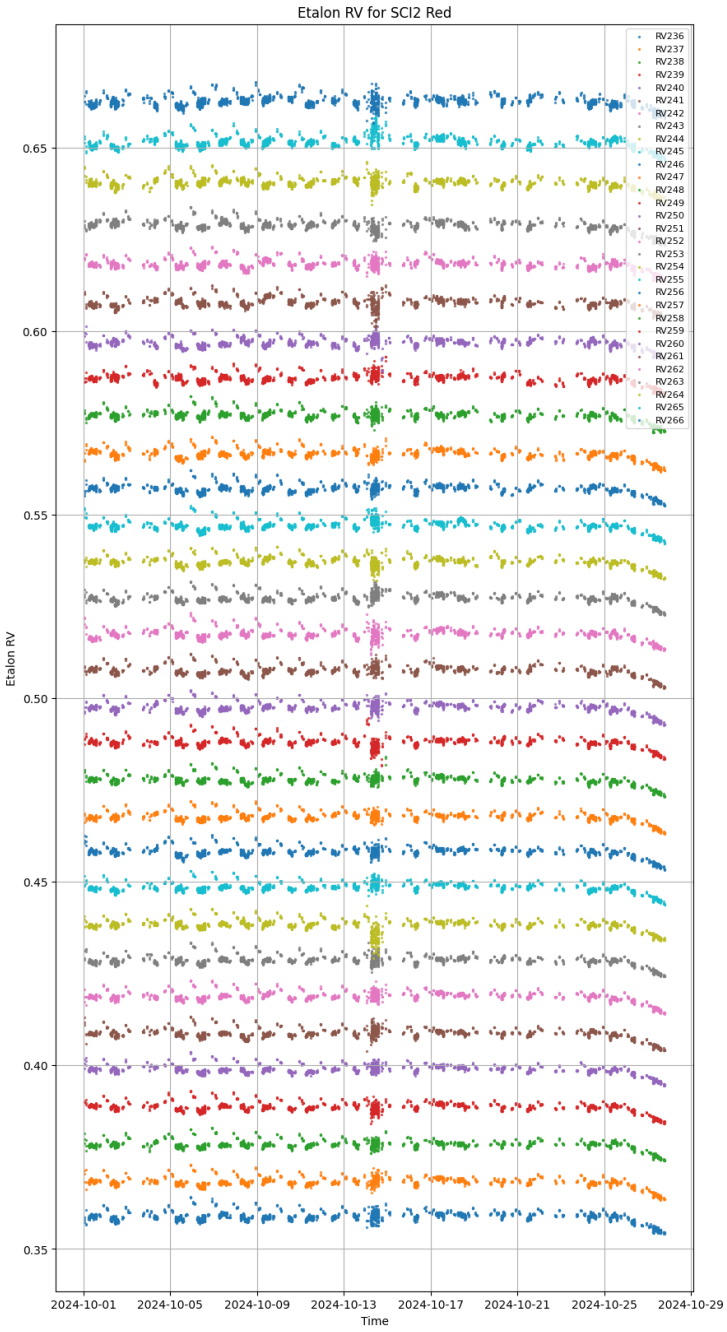

Figure 4: Red chip performance for SCI2 (orders 235-266). Most orders show consistent, flat behavior around a median value with small scatter.

Etalon RV for SCI3 Fiber, Green Chip

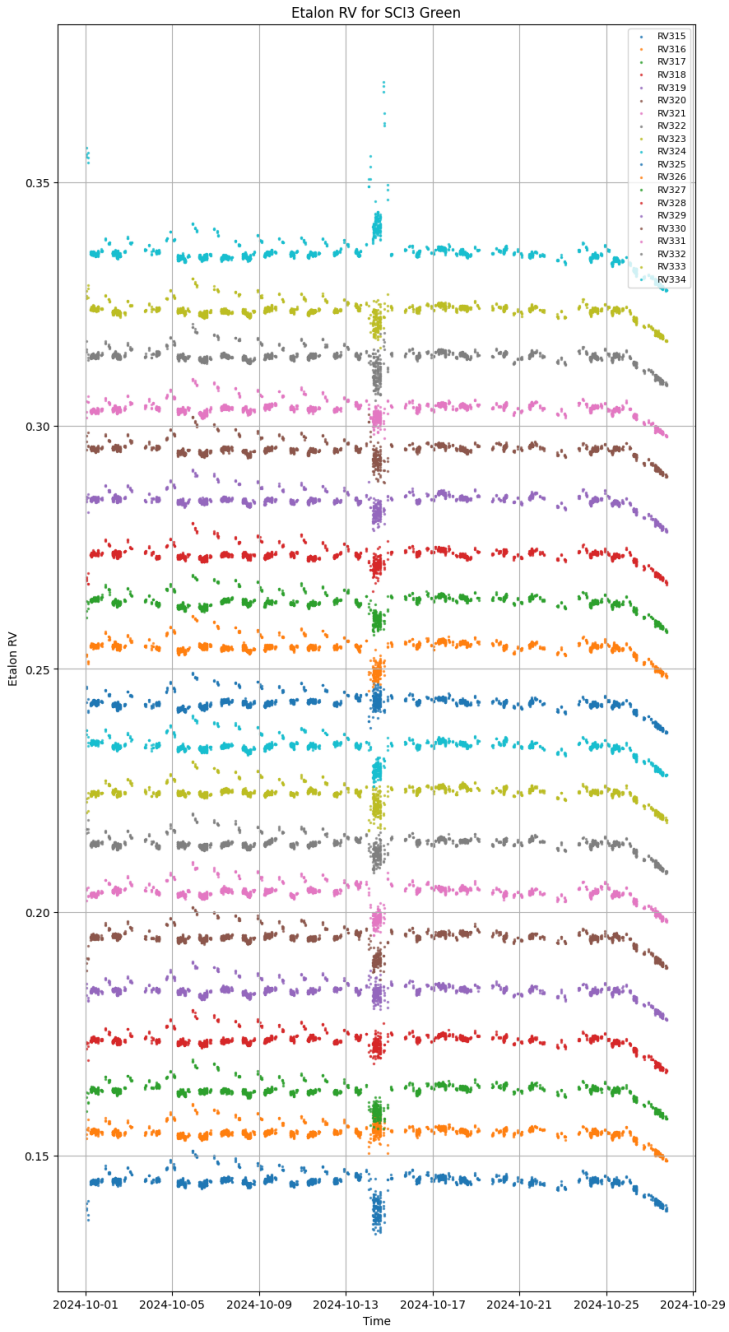

Figure 5: Green chip performance for SCI3 (orders 300-334). Most orders show consistent, flat behavior around a median value with small scatter.

Etalon RV for SCI3 Fiber, Red Chip

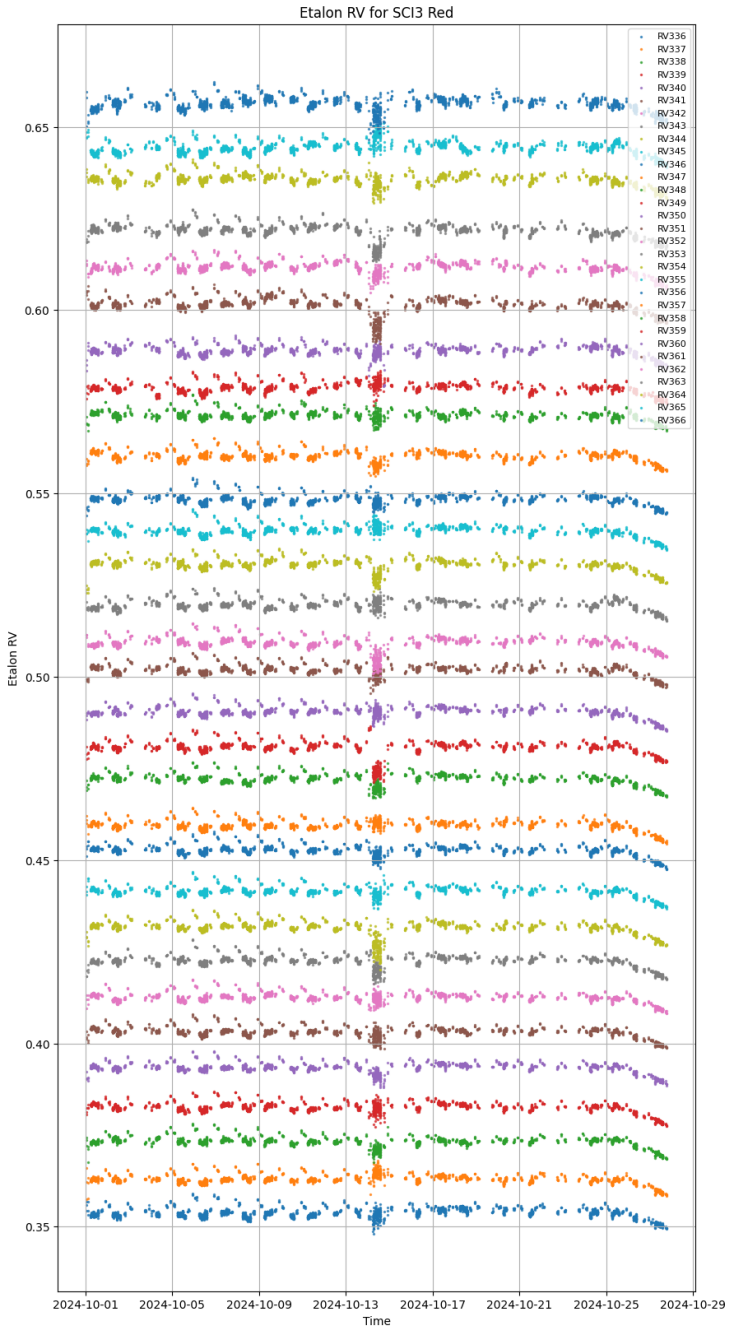

Figure 6: Red chip performance for SCI3 (orders 335-366). Most orders show consistent, flat behavior around a median value with small scatter.

Green SCI1: New DB vs. Old DB Comparison

A key comparison was performed between two versions of the database, focusing on the Green SCI1 etalon data. Refer to Figure 1 for the new database plot:

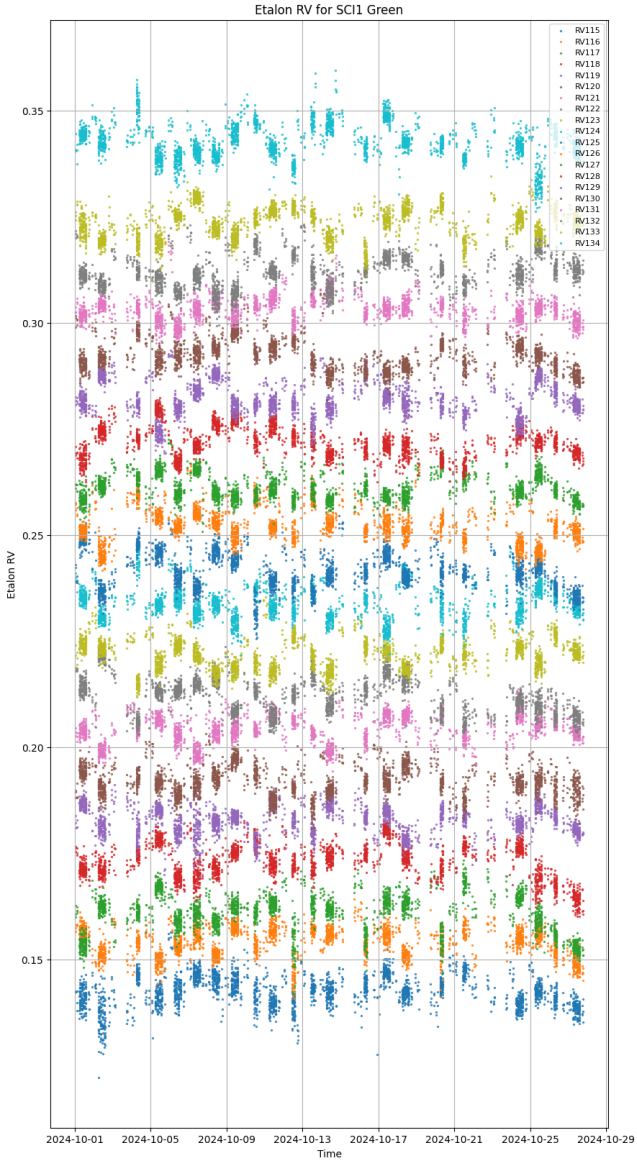

Figure 7: SCI1 green chip for the old database. There is more scatter apparent than in the new database,

- **Etalon mask cleanup:** The older version of the DRP used etalon masks that included many duplicate lines— repeated entries that were meant to represent the same etalon. These duplicates had slightly different positions, introducing artificial noise into the RV cross-correlation function. The new version of the DRP removes these duplicates, leading to cleaner CCFs and more stable RVs.

- **Wavelength solution handling:** The old DRP used interpolation between Laser Frequency Comb (LFC) calibrations to create a smoothed wavelength solution. While this suppressed drift, it also masked real instrumental changes. The new DRP disables interpolation. This makes any instrumental drift visible in the etalon RVs, allowing it to be corrected explicitly in future steps. As a side effect, the RVs now show more structure—but that structure reflects reality, not noise.

The resulting etalon RVs look smoother, as seen in the comparison between the old database and new database for SCI1 green chip orderlets. The new version will provide better inputs for physical modeling and correction.

Overall RMS Comparison: New vs. Old DB

New DB RMS Summary

	SCI Fiber	Max Order	Min Order	Max RMS(km/s rms)	Max Deviation(km/s rms + 1)	Min RMS(km/s rms)	Min Deviation(km/s rms + 1)
0	Green SCI1	34	16	0.003477	2.032442	0.001711	1.0
1	Green SCI2	21	16	0.003454	1.528524	0.002260	1.0
2	Green SCI3	25	30	0.007339	1.369381	0.005360	1.0
3	Red SCI1	31	21	0.008373	6.033667	0.001388	1.0
4	Red SCI2	22	29	0.003813	2.871218	0.001328	1.0
5	Red SCI3	21	31	0.011830	2.762873	0.004282	1.0

Figure 8: Summary of per-order RV scatter (RMS) for each fiber in the new database. For each fiber-chip combination, the table identifies the order with the highest RMS ("Max Order") and lowest RMS ("Min Order") along with their respective values. The Max and Min Deviation columns simply report RMS + 1. This table highlights the best- and worst-performing orders for each fiber, providing insight into intra-fiber consistency. The results reflect more accurate calibration. There was no consensus for the best and worst orderlets overall for all three fibers.

Old DB RMS Summary

	SCI Fiber	Max Order	Min Order	Max RMS(km/s rms)	Max Deviation(km/s rms + 1)	Min RMS(km/s rms)	Min Deviation(km/s rms + 1)
0	Green SCI1	17	34	0.010261	2.044750	0.005018	1.0
1	Green SCI2	17	34	0.008327	1.673068	0.004977	1.0
2	Green SCI3	17	30	0.011710	1.570001	0.007458	1.0
3	Red SCI1	31	24	0.011136	2.981164	0.003736	1.0
4	Red SCI2	31	5	0.004127	2.150399	0.001919	1.0
5	Red SCI3	22	29	0.013634	2.587432	0.005269	1.0

Figure 9: Summary of per-order RV scatter (RMS) for each fiber in the old database. As with the new DB, this table reports the order with the highest and lowest RMS for each fiber and includes the corresponding RMS values and RMS+1 deviations. Though the RMS values for Green SCI1 and SCI2 appear lower than in the new DB, these results are affected by duplicated mask lines and wavelength interpolation. In several cases, the worst-performing orders exhibit significantly elevated RMS, pointing to less consistent calibration in the old pipeline.

- **SCI1 Green:** In the old DB, RMS is artificially high. The time series plot, however, still shows inconsistent scatter. The new DB reveals more drift but is more coherent order-to-order and physically informative.
- **SCI2 and SCI3:** The new DB shows more uniform behavior across fibers, while the old DB had large RMS gaps due to calibration inconsistencies.
- **Red Chip:** The red chip remains noisier in both versions, but with the new DB, some of the most extreme outliers are eliminated thanks to mask improvements.

Supplemental Project: WLS File Consistency Checker

In parallel with the RV analysis, a separate utility script named `WLSChecker.py` was developed to validate internal metadata consistency in KPF FITS files. Specifically, the script checks whether the keywords `WLSFILE` and `WLSFILE2` contain matching date strings. Inconsistent values may indicate a data reduction error or misassociation of calibration files, which could affect scientific results.

Functionality

The script can be run in two modes:

- `-f /path/to/file.fits`: Checks a single FITS file.
- `-d /path/to/dir`: Recursively checks all FITS files in a directory tree.

If a mismatch is detected:

- A new FITS header keyword `WLSMATCH` is written with value `"no"`.

- The script can optionally list only the mismatches using `-l`.
- A `-st` option allows automatic triggering of `kpf_slowtouch.sh` to retouch affected files and flag them for reprocessing.

Implementation

The script uses `astropy.io.fits` to read and update FITS headers, and regular expressions to extract dates from `WLSFILE` and `WLSFILE2`. The logic ensures robust handling of missing keywords and non-date formats, with structured logging and graceful error handling.

Sample output:

```
Processed KP.20241015.12024.27_L2.fits: WLSFILE=20240912, WLSFILE2=20241015, WLSMATCH='no'
Calling kpf_slowtouch.sh on: KP.20241015.12024.27.fits
```

Conclusion

This project investigated discrepancies in radial velocity (RV) stability across individual spectral orders for each science fiber and chip using etalon calibration data from the Keck Planet Finder. By plotting order-by-order RV time series and getting RMS per orderlet, we identified which fibers and orders consistently perform better or worse in terms of calibration stability.

This per-orderlet, per-fiber perspective adds critical insight that would be lost in bulk metrics like overall fiber RMS or chip-level stability. It helps pinpoint which orders may need to be excluded, down-weighted, or further investigated in future calibration models.