CrossMark

The Extreme Stellar-signals Project. III. Combining Solar Data from HARPS, HARPS-N, EXPRES, and NEID

Lily L. Zhao[1] , Xavier Dumusque[2] , Eric B. Ford[3,4,5] , Joe Llama[6] , Annelies Mortier[7] , Megan Bedell[1] ,
Khaled Al Moulla[2] , Chad F. Bender[8] , Cullen H. Blake[9] , John M. Brewer[10] , Andrew Collier Cameron[11,12] ,
Rosario Cosentino[13] , Pedro Figueira[2,14] , Debra A. Fischer[15] , Adriano Ghedina[13] , Manuel Gonzalez[13],
Samuel Halverson[16] , Shubham Kanodia[3,4,17] , David W. Latham[18] , Andrea S. J. Lin[3,4] , Gaspare Lo Curto[19] ,
Marcello Lodi[13] , Sarah E. Logsdon[20] , Christophe Lovis[2] , Suvrath Mahadevan[3,4,21] , Andrew Monson[8] ,
Joe P. Ninan[22] , Francesco Pepe[2] , Rachael M. Roettenbacher[23] , Arpita Roy[24,25] , Nuno C. Santos[14,26] ,
Christian Schwab[27] , Guðmundur Stefánsson[28,33] , Andrew E. Szymkowiak[15,29] , Ryan C. Terrien[30] , Stephane Udry[2] ,
Sam A. Weiss[15] , François Wildi[2] , Thibault Wildi[31] , and Jason T. Wright[3,4,32]

[1] Center for Computational Astrophysics, Flatiron Institute, Simons Foundation, 162 Fifth Avenue, New York, NY 10010, USA; lzhao@flatironinstitute.org
[2] Astronomy Department of the University of Geneva, 51 ch. des Maillettes, 1290 Versoix, Switzerland
[3] Department of Astronomy & Astrophysics, 525 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA
[4] Center for Exoplanets and Habitable Worlds, 525 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA
[5] Institute for Computational & Data Sciences, The Pennsylvania State University, University Park, PA 16802, USA
[6] Lowell Observatory, 1400 W. Mars Hill Road, Flagstaff, AZ 86001, USA
[7] School of Physics & Astronomy, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK
[8] Steward Observatory, University of Arizona, 933 N. Cherry Avenue, Tucson, AZ 85721, USA
[9] Department of Physics and Astronomy, University of Pennsylvania, 209 S 33rd Street, Philadelphia, PA 19104, USA
[10] Dept. of Physics & Astronomy, San Francisco State University, 1600 Holloway Avenue, San Francisco, CA 94132, USA
[11] SUPA, School of Physics & Astronomy, University of St Andrews, North Haugh, St Andrews, KY169SS, UK
[12] Centre for Exoplanet Science, University of St Andrews, North Haugh, St Andrews, KY169SS, UK
[13] Fundación Galileo Galilei-INAF, Rambla José Ana Fernandez Pérez 7, E-38712 Breña Baja, Tenerife, Spain
[14] Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, 4150-762 Porto, Portugal
[15] Department of Astronomy, Yale University, 52 Hillhouse Avenue, New Haven, CT 06511, USA
[16] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[17] Earth and Planets Laboratory, Carnegie Institution for Science, 5241 Broad Branch Road, NW, Washington, DC 20015, USA
[18] Center for Astrophysics, Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[19] European Southern Observatory, Av. Alonso de Cordova 3107, Casilla 19001, Santiago de Chile, Chile
[20] NSF's National Optical-Infrared Astronomy Research Laboratory, 950 N. Cherry Avenue, Tucson, AZ 85719, USA
[21] ETH Zurich, Institute for Particle Physics & Astrophysics, Zurich, Switzerland
[22] Department of Astronomy and Astrophysics, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai 400005, India
[23] Department of Astronomy, University of Michigan, 1085 S. University Ave., Ann Arbor, MI 48109, USA
[24] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[25] Department of Physics and Astronomy, Johns Hopkins University, 3400 N Charles Street, Baltimore, MD 21218, USA
[26] Departamento de Física e Astronomia, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 4169-007 Porto, Portugal
[27] School of Mathematical and Physical Sciences, Macquarie University, Balaclava Road, North Ryde, NSW 2109, Australia
[28] Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08540, USA
[29] Department of Physics, Yale University, 217 Prospect Street, New Haven, CT 06511, USA
[30] Carleton College, One North College Street, Northfield, MN 55057, USA
[31] Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, D-22607 Hamburg, Germany
[32] Penn State Extraterrestrial Intelligence Center, 525 Davey Laboratory, The Pennsylvania State University, University Park, PA 16802, USA

Abstract

We present an analysis of Sun-as-a-star observations from four different high-resolution, stabilized spectrographs —HARPS, HARPS-N, EXPRES, and NEID. With simultaneous observations of the Sun from four different instruments, we are able to gain insight into the radial velocity precision and accuracy delivered by each of these instruments and isolate instrumental systematics that differ from true astrophysical signals. With solar observations, we can completely characterize the expected Doppler shift contributed by orbiting Solar System bodies and remove them. This results in a data set with measured velocity variations that purely trace flows on the solar surface. Direct comparisons of the radial velocities measured by each instrument show remarkable agreement with residual intraday scatter of only 15–$30\,\mathrm{cm\,s^{-1}}$. This shows that current ultra-stabilized instruments have broken through to a new level of measurement precision that reveals stellar variability with high fidelity and detail. We end by discussing how radial velocities from different instruments can be combined to provide powerful leverage for testing techniques to mitigate stellar signals.

---

[33] NASA Sagan Fellow.

1. Introduction

The radial velocity (RV) method of discovering exoplanets measures the center-of-mass motion of a planet's host star as both the planet and star orbit their common center of mass. This center-of-mass motion induces a periodic Doppler shift on the star's spectrum over time. Highly stabilized spectrographs are able to measure this Doppler shift down to submeter-per-second levels (Pepe et al. 2004; Cosentino et al. 2012; Pepe et al. 2013; Jurgenson et al. 2016; Schwab et al. 2016).

Changes on a star's surface will introduce spectral variations that can be mistakenly measured as a true center-of-mass shift of the spectra (Meunier 2021). These surface variations include, but are not limited to, p-mode oscillations (Mayor et al. 2003; Bouchy et al. 2005; Kjeldsen et al. 2005; Arentoft et al. 2008; Chaplin et al. 2019), granulation (Dravins 1982; Kjeldsen & Bedding 1995; Lindegren & Dravins 2003; Dumusque et al. 2011b; Meunier et al. 2015; Cegla et al. 2018; Lanza et al. 2019), and supergranulation (Rieutord & Rincon 2010; Rincon & Rieutord 2018; Meunier & Lagrange 2019). Activity features—such as spots, faculae, and plages (Saar & Donahue 1997; Hatzes 2002; Saar 2003; Desort et al. 2007; Huélamo et al. 2008; Boisse et al. 2011; Dumusque et al. 2011a; Lovis et al. 2011; Jeffers et al. 2013; Santos et al. 2014; Cabot et al. 2021; Roettenbacher et al. 2021)—can also introduce periodic spectral variations as they rotate in and out of view as well as give rise to nonperiodic signals as these activity regions and their properties evolve with time.

Measured RV variations due to stellar surface variations contribute scatter or potentially periodic offsets to RV time series (Crass et al. 2021). Different types of variations will manifest in the spectra in different ways, in different lines, and on different timescales. Spectral lines affected by these surface variations may be shallower/deeper, appear shifted, or acquire asymmetric wings. Some spectral lines are likely to be more affected than others, depending on the line's wavelength, species, and formation temperature (Dumusque 2018; Wise et al. 2018; Ning et al. 2019; Cretignier et al. 2020; Al Moulla et al. 2022). These variations will introduce (potentially periodic) RV scatter of an amplitude far greater than the $\sim 10 \, \mathrm{cm \, s^{-1}}$ expected signal for an Earth-like planet orbiting a Sun-like star.

Many methods have been developed to mitigate the variations that arise in RV measurements due to stellar signals (e.g., Aigrain et al. 2012; Jones et al. 2017; Gilbertson et al. 2020; Lafarga et al. 2020; Rajpaul et al. 2020; Zhao & Tinney 2020; Collier Cameron et al. 2021; Cretignier et al. 2021). The previous paper in this series (The EXPRES Stellar Signals Project II;[34] Zhao et al. 2022) performed a head-to-head comparison of different mitigation methods tested on the same set of real observations for four different stars. This direct comparison produced a confusing outcome—methods often disagreed on the signals that should be attributed to stellar surface variations, even though all methods used identical data sets.

That exercise highlighted the difficulty of rigorously establishing method performance. The choice to test methodologies using real versus simulated data sets highlights different strengths and weaknesses of the methods and introduces very different working hypotheses. Simulated data sets come with the benefit of knowing the "answer at the back of the book"—in other words, the resulting data set is one for which the structure of the injected signals are known. It is also known whether there are any true center-of-mass shifts (simulated planetary signals) that mitigation methods need to preserve. As with all simulations, however, it is unclear whether the simulated signals are physically motivated and complete enough to be representative of a real-case scenario. Using real data sets, however, introduces nonastrophysical sources of systematics (e.g., instrumental variation, different observing conditions/limitations, etc.) and the potential of undetected planetary signals, which can lead to confusion and misinterpretations of results.

Sun-as-a-star observations—i.e., disk-integrated observations of the Sun as if it were a star—have many of the benefits of both simulated data and real observations. Because they are observations of a real star, solar observations are sure to capture realistic manifestations of stellar variability. True center-of-mass shifts are well-understood for the Sun, as Solar System bodies are characterized well beyond the limits of current RV capabilities. Observing the Sun is also relatively inexpensive, as observations occur during the day when the spectrographs are not otherwise used. With such a bright source, it is easy to achieve observations with high signal-to-noise ratios (S/Ns) as well as temporal sampling on a timescale of minutes.

Several planet-hunting spectrographs therefore have solar feeds—i.e., the High-Accuracy Radial-velocity Planet Searcher (HARPS), the High-Accuracy Radial-velocity Planet Searcher in the North (HARPS-N; Phillips et al. 2016), the EXtreme PREcision SPectrograph (EXPRES; J. Llama 2023, in preparation), the NN-explore Exoplanet Investigations with Doppler spectroscopy instrument (NEID; Lin et al. 2022), and the Keck Planet Finder (KPF; R. Rubenzahl 2023, in preparation). Other instruments have plans to add a solar feed—e.g., on the Échelle Spectrograph for Rocky Exoplanet and Stable Spectroscopic Observations (ESPRESSO; Leite et al. 2022) or the M dwarf Advanced Radial velocity Observer Of Neighboring eXoplanets (MAROON-X).

Solar observations also feature some unique disadvantages when compared with nighttime stellar observations. As a resolved source, solar observations must contend with differences in observing conditions (e.g., cloud coverage, airmass, etc.) across the disk of the Sun. The Earth's relative velocity with respect to the Sun is lower than the relative velocity of the Earth to other stars, meaning telluric lines, from absorption by Earth's atmosphere, do not shift as much relative to solar lines as stellar lines. Observations that span a variety of stellar types and stellar parameters beyond the Sun will ultimately be necessary to more holistically understand how surface variations present in stabilized spectral data. However, acquiring solar data is a valuable starting point for understanding stellar signals as measured by precise radial-velocity (PRV) spectrographs.

---

[34] This is the previous name of this series. With this installment, the series name is changed to "The Extreme Stellar-signals Project" as the analysis expands beyond EXPRES data.

We combine data from four different PRV spectrographs—HARPS (Pepe et al. 2004), HARPS-N (Cosentino et al. 2012), EXPRES (Jurgenson et al. 2016), and NEID (Schwab et al. 2016). Having concurrent data of the same source allows us to benchmark each instrument against the other three. This can reveal how differences in instrument design, instrument systematics, and data reduction processes might affect how the same stellar signals manifest in the data collected by each instrument.

Here, we present and discuss this combined solar data set. In Section 2, we describe the different instruments and their corresponding data sets, and we quantify the degree to which they overlap with one another. We also discuss how each instrument's pipeline contends with some of the disadvantages of observing the Sun. Section 3 explains how the RVs from the different instruments are compared across different timescales; results from this comparison are outlined in Section 4. We end with conclusions and a discussion of future directions in Section 5.

2. Data

We combine data from HARPS, HARPS-N, EXPRES, and NEID. We use observations collected between 2021 May 25 and June 23, which were selected as dates that contained a large number of overlapping data from the four instruments. Only exposures taken at an airmass less than 2.2 are used. Including observations taken at an airmass greater than 2.2 introduces significant error from atmospheric extinction. With the Sun, higher airmasses also increase the differential extinction effect of the resolved disk of the Sun. However, enacting an airmass cut that is harsher would significantly reduce the amount of overlapping data we have between different instruments.

The observations from each instrument are plotted in Figure 1 as black points. Colored points show the observed RVs from each instrument binned to five-minute intervals (see Section 3.1 for a detailed description of the binning process). We will only use "observations" to refer to the RV measurements returned by each instrument's data reduction pipeline (i.e., the binned RVs will be referred to as "binned RVs" or "binned points" rather than observations).

This analysis makes use of the RVs from each instrument's default data reduction pipeline (DRP). In other words, the RVs used are produced by the same pipeline that is used to extract the nightly data for facility users of each instrument with some additional Sun-specific corrections (e.g., implementing bary-centric corrections as relevant to solar observations, accounting for the Sun being a resolved source, and others discussed further below). The RVs used here have had known Solar System planet signals removed using the known ephemerides of these objects as curated by the Jet Propulsion Laboratory (JPL; Park et al. 2021). These shifts were removed at the wavelength level for each spectrum. For more details about each instrument's DRP, see Dumusque et al. (2021) for HARPS and HARPS-N, Petersburg et al. (2020) for EXPRES, and https://neid.ipac.caltech.edu/docs/NEID-DRP/ for NEID.

Figure 2 shows a two-hour subset of the observed RVs taken on 2021 May 31 from each instrument (i.e., the figure shows the observed RVs as provided by each instrument's pipeline, not the RVs binned to shared time stamps). As it is impractical to show all the data at this high time resolution, we show here the period of time with the greatest amount of overlap between
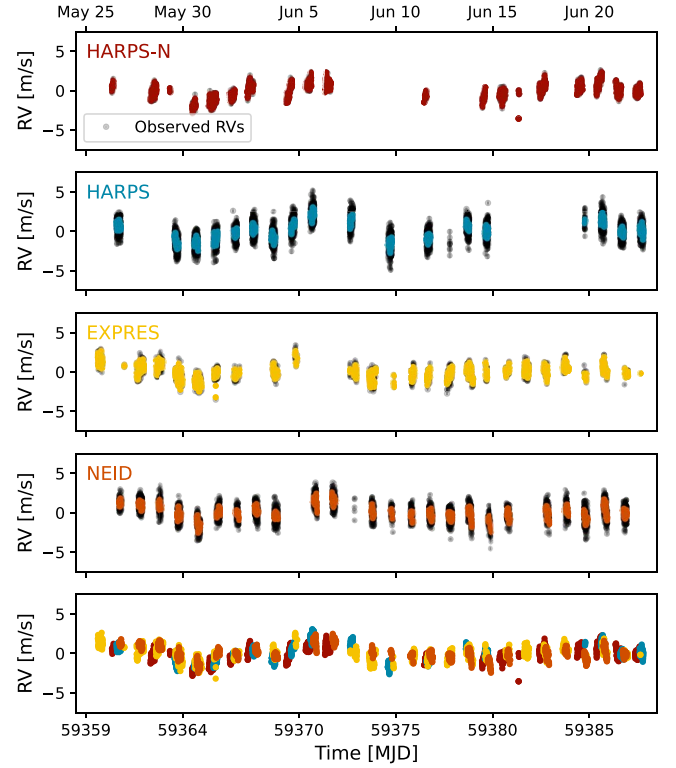

Figure 1. Solar RVs from HARPS-N, HARPS, EXPRES, and NEID. Observed RVs are shown as black points; colored points represent the RVs binned to shared time stamps with an interval of five minutes throughout a day. The color assigned to the data for each instrument is consistent throughout this work. The bottom most subplot overlays the binned RVs from all instruments (with a separate offset applied to each instrument). All instruments trace similar long-term RV variations.
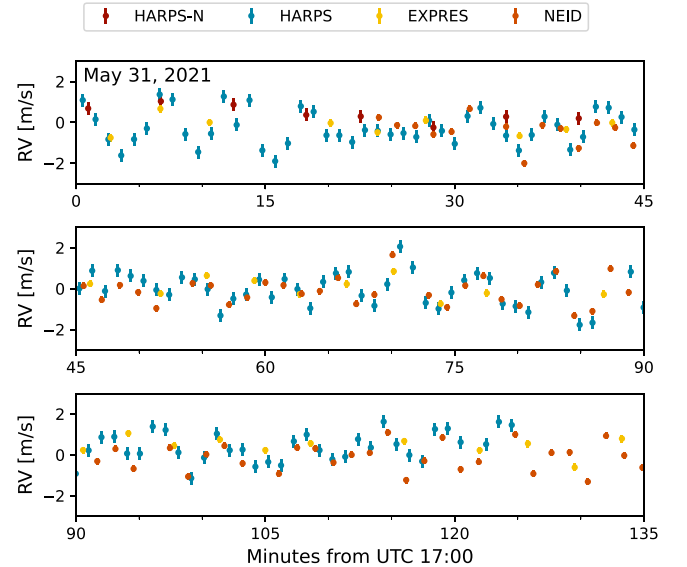

Figure 2. Observed RVs from HARPS-N, HARPS, EXPRES, and NEID. DRP RVs (i.e., not binned RVs) are shown over two hours on 2021 May 31 from UTC 17:00 to UTC 19:15. A separate RV offset, calculated as the median of all RVs from that instrument taken on the day, has been applied to each instrument. The x-axis is given in minutes from UTC 17:00 with minor tick marks every 5 minutes (which is close to the ν_{max} for the Sun). All four instruments trace similar short-term RV variations.

Table 1
Instrument/Data Properties

Parameter	HARPS	HARPS-N	EXPRES	NEID
Instrument				
Solar Telescope Name	HELIOS	LCST	LOST	NEID Solar Feed
Location	La Silla, Chile	La Palma, Spain	Flagstaff, AZ, USA	Kitt Peak, AZ, USA
Latitude	29°15′27″ S	28°45′49″ N	34°44′40″ N	31°57′30″ N
Longitude	70°44′15″ W	17°53′41″ W	111°25′19″ W	111°35′48″ W
Start of Spectrograph Science Operations	2003	2012	2019	2021
Start of Solar Observations	2018	2015	2020	2020
Full Wavelength Range [nm]	378–691	383–690	390–780	380–1046
Median Resolution	115,000	118,000	137,000	120,000
Solar Telescope Components				
Aperture of Solar Telescope Lens [mm]	75	75	75	75
Lens Coating	MgF_2	MgF_2	MgF_2	VIS-NIR[a]
Solar Telescope Mount	Alt/Az	Alt/Az	Equatorial	Alt/Az
Dome Material	Polymethyl Methacrylate	Polymethyl Methacrylate	Polymethyl Methacrylate	None
Fiber Feed Length [m]	∼30	∼20	∼80	∼45
Data				
Amount of Data[b]				
Total Observations	4570	1082	1459	3617
Binned RVs	849	1176	1218	1059
Days w/ Data (out of 29)	19	19	26	25
Avg. Obs./Day	157	37	50	124
Cadence				
Exposure Time [s]	30	300	178[c]	55
Dead Time [s]		26	52	28
Data Quality				
Average RV Error [m s^{-1}]	0.3	0.29	0.19	0.16
Average Binned RV Error [m s^{-1}]	0.09	0.21	0.12	0.06
S/N at 4200 Å	140	130	100	250
S/N at 4750 Å	280	220	270	420
S/N at 5500 Å	400	290	480	570
S/N at 6250 Å	420	310	620	630
S/N at 6650 Å	410	300	660	640
Start/Stop Criteria[d]				
Start Airmass	5.1	2.9	2.7	*∼1.3*
Stop Airmass	4.6	*∼1.3*	*∼2.0*	*∼1.3*
Start Time [UTC]	*∼12:37*	*∼7:55*	*∼13:11*	16:31
Stop Time [UTC]	*∼20:40*	16:00	(+1) 00:00	22:30

Notes.
[a] The NEID lens makes use of the proprietary VIS-NIR antireflection coating from Edmund Optics (Lin et al. 2022).
[b] Specifically the amount of data used in this analysis, which only includes data collected between 2021 May 25 and June 23.
[c] An average exposure time is given here because EXPRES uses a set S/N rather than a fixed exposure time (Petersburg et al. 2018).
[d] Italicized values indicate that those values were calculated as opposed to the values actually used to start/stop observations.

the four instruments. This is also the only time span within the month for which there are good-quality solar observations from all four instruments (from ∼17.4 to 17.7 UTC time of day). Figure 2, therefore, shows the optimal time frame for a direct comparison of the DRP RVs from all four instruments on short timescales. Comparisons that take into account all times when instruments overlap across the entire month are done using the binned RVs and are discussed more below.

Table 1 gives instrument and data properties for each data set. For each instrument, we list the start of science operations (i.e., the time at which each instrument is considered to have been fully commissioned), the start of regular observations for each instrument's corresponding solar telescope, the full wavelength range of instrument, and the median resolution across each instrument's detector. We also compare a subset of hardware components across solar telescopes that differ between instruments. All four solar telescopes employ a

75 mm achromatic doublet lens from Edmund Optics (NEID has a slightly different coating to address its redder wavelength range), and the light is combined via a 2″ Polytetrafluoroethylene (PTFE) integrating sphere from Thorlabs. EXPRES is the only telescope of the four to use an equatorial mount. HARPS, HARPS-N, and EXPRES all use the same type of telescope dome (NEID's solar telescope has no dome).

To capture the number and cadence of observations, Table 1 gives the amount of data, exposure time, and average readout time for each instrument. For each data set, we give both the total number of observations and the total number of binned data points. The number of days (out of the 29 included in this analysis) on which there is data from each instrument is also given. We note that EXPRES terminates exposures at a set S/N rather than at a fixed time; therefore, an average exposure time is cited for EXPRES in Table 1.
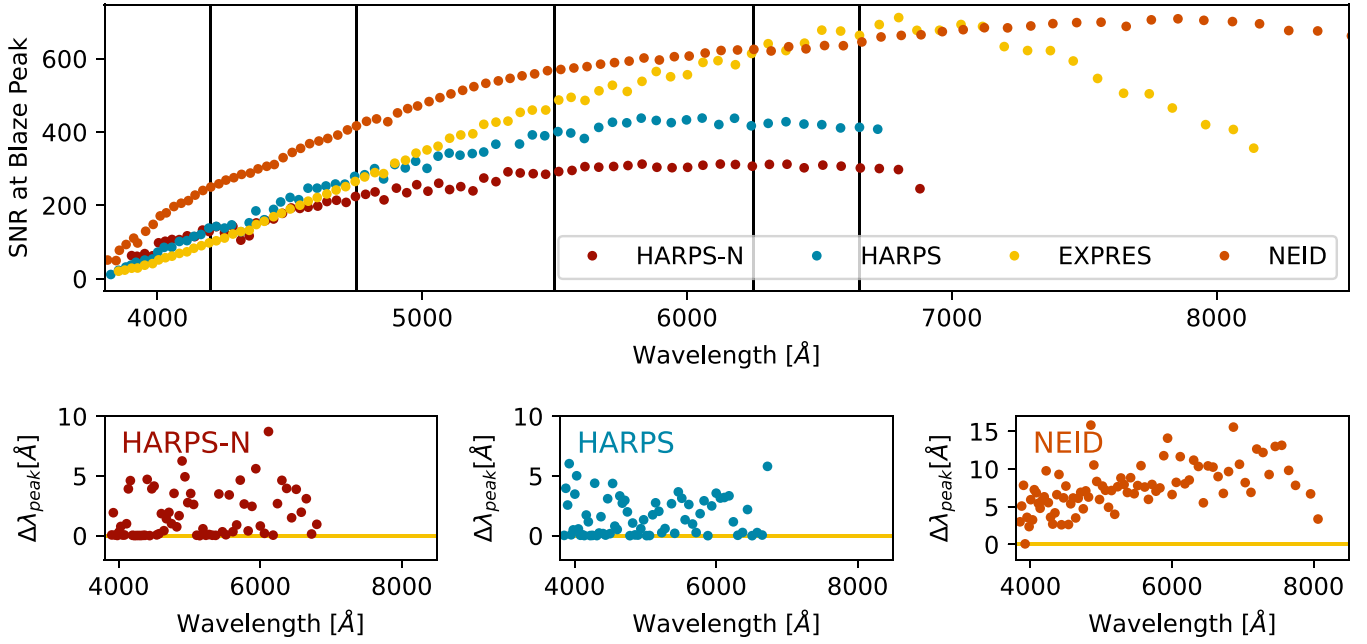
Figure 3. Top row: per-exposure solar S/N at the blaze peak of every échelle order for a representative observation from each instrument. Black vertical lines mark the wavelengths for which the average S/N is reported in Table 1. We note that NEID's wavelength range extends further red beyond the edge of the plot. Bottom row: difference between the wavelength of an instrument's blaze peak, λ_{peak}, for each order. From left to right, HARPS-N, HARPS, and NEID are compared to the blaze peak wavelength of EXPRES. For HARPS-N and HARPS, the λ_{peak} of each order falls on average within 2 Å of the EXPRES λ_{peak}; between NEID and EXPRES, the λ_{peak} is on average within 7 Å.

The average quality of observations from each instrument is given in terms of the average analytic RV error and the S/N (per 1D extracted pixel) of a representative observation for a range of wavelengths. The average analytic RV error given in Table 1 is directly calculated using just the month of shared data. The quoted value may therefore differ from the approximate analytic errors expected from each instrument's DRP, which is typically calculated using a larger and more varied set of observations. For HARPS and HARPS-N, the S/N per 1D extracted pixel is calculated from the flux of each pixel divided by the expected photon noise and detector noise added in quadrature. In the case of high-S/N observations, like those of the Sun, the S/N is close to the square root of the flux. For EXPRES, the S/N is directly calculated as the square root of the photon count. For NEID, the S/N is calculated from the ratio of the science fiber's flux and the science fiber's analytic variance, which is given by the error associated with flat-relative optimal extraction as described in Zechmeister et al. (2014). Similarly to the other instruments, the analytic variance calculated by NEID's DRP is dominated by photon noise, especially at high S/N.

Figure 3 shows the average peak S/N within each échelle order—i.e., the peak of the blaze function within that order—plotted against the wavelength at which the peak occurs, λ_{peak}. The average S/N at five wavelengths, as marked in Figure 3 by vertical black lines, is listed in Table 1. The bottom row of subplots in Figure 3 shows the difference in λ_{peak} for HARPS-N, HARPS, and NEID as compared to the λ_{peak} of EXPRES in each order. Between the different instruments, the position of the blaze peak differs on average by less than 2 Å between EXPRES and both HARPS-N and HARPS, which corresponds to less than 1.5%–3% of the order width. The blaze peak for NEID differs from EXPRES by less than 7 Å on average, or 7.5% of the NEID order width. This agreement is expected

because all instruments make use of an R4 échelle grating with 31.6 lines per millimeter.

Table 1 ends with the start/stop criteria used by each instrument, i.e., when solar observations are scheduled to begin and end for each day. NEID uses time of day as its start/stop criteria, and observations are fully automated in the instrument computer. HARPS and HARPS-N uses the altitude of the Sun to guide the beginning and end of solar observations. Observations are manually started by the telescope's day operators, and observations proceed until manually stopped for late-afternoon calibrations. EXPRES is fully automated and uses the altitude of the Sun to start operations and time of day to stop observing. Instruments using altitude-based stop criteria are often stopped earlier if the instrument is needed to prepare for nighttime observing. Start/stop times are given in UTC time of day, not in the local time of the telescope location.

We give the observing window criteria used by each instrument in the table as roman (i.e., nonitalic) text. We also compute the corresponding value (i.e., the airmass that corresponds to a start/stop time or the start/stop time that corresponds to an airmass) to allow for easier comparison across all instruments. These calculated values are presented in the table in italicized text. The conversions are done for a day in the middle of the shared time series (2021 June 8).

The timing of the solar observations from each instrument is diagrammed in Figure 4. The top of Figure 4 shows the airmass of the Sun in June across a day as seen by each instrument. This is shown as a function of UTC time of day. The bottom plot of the figure shows the time stamps of each observation over the 29 days of data used in the analysis. The day as given by its MJD (left y-axis) or calendar day (right y-axis) increases from top to bottom. Comparing the time stamps of the observations shows when and how often observations from different instruments overlap, as well as gaps in the overlap time. The
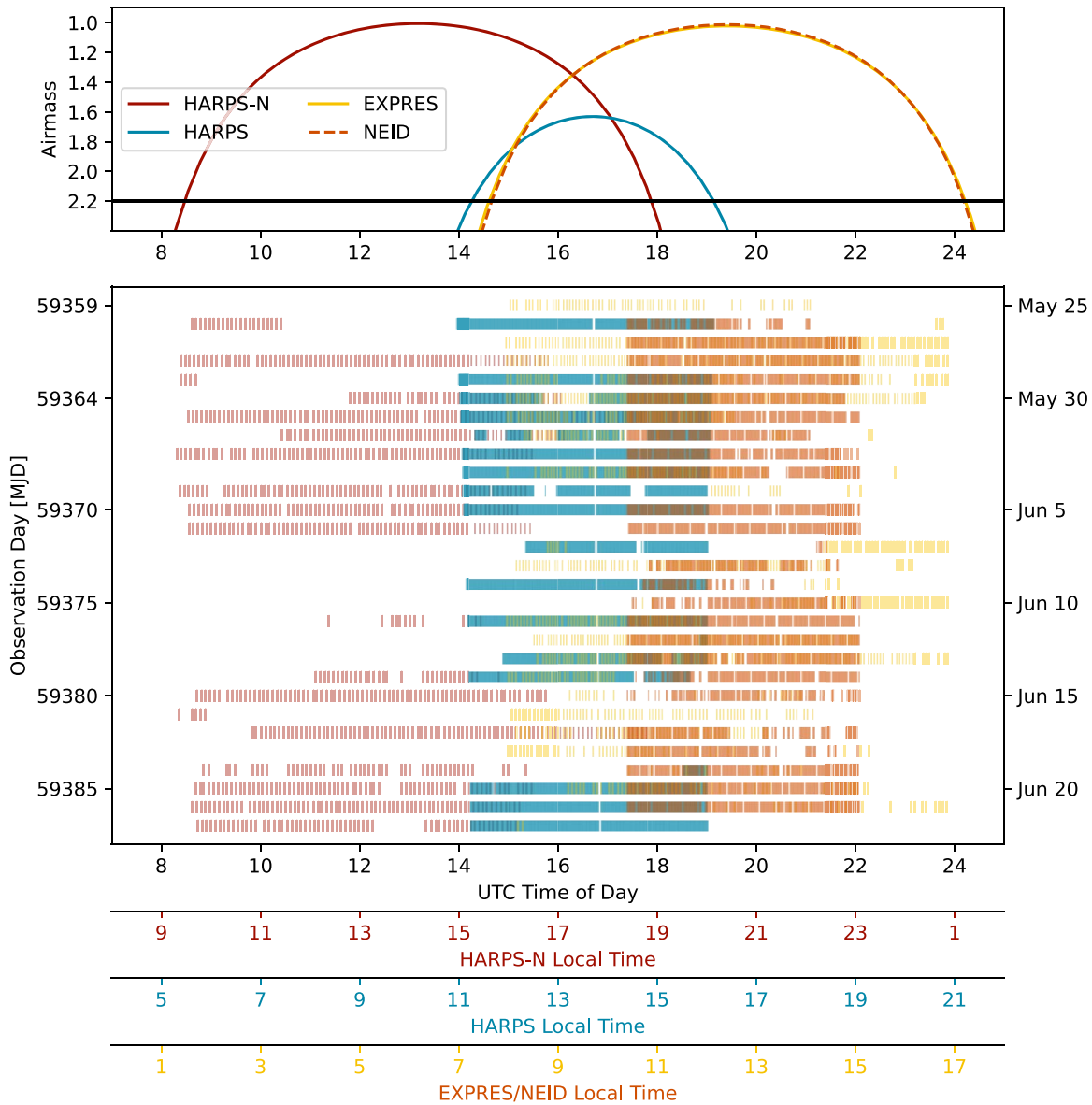
Figure 4. The time stamps of observations from each instrument as a function of time of day. Top: airmass of the Sun in June as a function of UTC time of day as seen from each of the four instruments. The EXPRES and NEID curves largely overlap. The analysis presented here uses only data taken at an airmass lower than 2.2 (marked by the horizontal black line). Bottom: Time stamps of observations from each instrument for each day (y-axis, increasing top to bottom) and UTC time of day (x-axis). This layout highlights the cadence of each time series as well as the times at which observations from different instruments overlap or when there are gaps in the observations. To help interpret UTC time of day, the local time at each instrument's location is given on separate axes below this plot. (Because EXPRES and NEID are located in the same time zone, they share a local time axis that is therefore colored yellow and orange.)

corresponding local time for each instrument is given in the additional x-axes at the bottom of the plot.

2.1. HARPS

The High Accuracy Radial Velocity Planet Searcher (HARPS) is a stabilized, fiber-fed, optical (378–691 nm) échelle spectrograph (Pepe et al. 2004). It was fully commissioned at the 3.6 m telescope at La Silla Observatory, Chile in 2003 October. HARPS has a median resolution of $R \sim 115,000$ and an instrument calibration precision of $50 \, \mathrm{cm \, s^{-1}}$. Daily solar observations show an RV rms that is below $50 \, \mathrm{cm \, s^{-1}}$

(where data have been binned over five minutes to average over the p-mode oscillation signal), demonstrating the short-term stability of this instrument.[35]

HARPS has also demonstrated long-term stability close to $1 \, \mathrm{m \, s^{-1}}$. Cretignier et al. (2021) shows that observations of Tau Ceti over 13 yr give an RV rms of $1.18 \, \mathrm{m \, s^{-1}}$, where the RVs delivered by the official HARPS pipeline have been daily binned to mitigate stellar oscillation and granulation signals. The RV rms of the same data set can be reduced to $1.02 \, \mathrm{m \, s^{-1}}$ with the use of post-processing techniques, such as *YARARA*, to mitigate instrumental systematics that are challenging to correct

---

[35] This is the case when using the implementation of the ESPRESSO pipeline for HARPS data (https://www.eso.org/sci/software/pipelines/espresso/espresso-pipe-recipes.html), as done in this paper for the Sun. It should be noted that this pipeline is still in development.

for at the extraction level (Cretignier et al. 2021). Cretignier (2023) shows that an RV rms of smaller than 1 m s^{-1} can be achieved for ~10 yr of observations of four other HARPS RV standards, and that new planetary candidates around these stars with amplitudes as small as 0.5 m s^{-1} for periods as long as 600 days can be detected.

The HARPS Experiment for Light Integrated Over the Sun (HELIOS) is a 75 mm telescope with a fiber feed into HARPS. It has been observing the Sun as a star since first light in 2018. HELIOS starts observing the Sun when it rises above an altitude of 10° and is programmed to stop when the Sun sets below 10°. In practice, solar observations are typically stopped much earlier, to prepare for nighttime observations. Exposure times are fixed at 30 s, for which HARPS reaches an analytical error per observation of 30 cm s^{-1}.

The pipeline for the HELIOS data is similar to the pipeline for HARPS-N, as described in Dumusque et al. (2021). It is largely the same pipeline as used for the nightly stellar data observed with either ESPRESSO or HARPS-N. HELIOS observes through a plexiglass dome and thus observes even when it is cloudy (although observations are stopped when the guiding camera cannot find the solar disk). Data quality is therefore assessed a posteriori via a Bayesian mixture-model approach, which takes in the S/N and airmass of the data for a given day, to vet observations during which the Sun was partially obscured by clouds and/or other weather conditions (Collier Cameron et al. 2019).

This procedure also produces a daily atmospheric extinction coefficient. This value is subsequently used to correct the solar RVs for differential atmospheric extinction effects that arise because the Sun is a resolved disk in the sky. Light from the lowest point of the Sun therefore travels through more of the atmosphere and has a higher airmass than light from the top of the Sun; this creates a daily downward trend in the RVs. Details of this extinction correction and the Bayesian mixture model as applied to HARPS and HARPS-N are described further in Al Moulla et al. (2023) and Collier Cameron et al. (2019), respectively.

2.2. HARPS-N

The High Accuracy Radial Velocity Planet Searcher in the North (HARPS-N) is a stabilized, fiber-fed, optical (383–690 nm) échelle spectrograph (Cosentino et al. 2012). It was fully commissioned at the 3.6 m Telescopio Nazionale Galileo (TNG) at La Palma, Spain in 2012 August. HARPS-N has a median resolution of $R \sim 118,000$ and instrument calibration precision of 49 cm s^{-1} (as measured by the median absolute offset between consecutive wavelength solutions). The on-sky performances are very similar to what is obtained with HARPS, with a daily RV rms for solar observation (5 minutes integration time) at around 40 cm s^{-1}.

HARPS-N exhibits similar long-term stability as HARPS. John et al. (2023) show that RV standard star HD 144579 returns an RV rms of 1.29 m s^{-1} over 9 yr of observations binned nightly, similar to the precision reached by HARPS for Tau Ceti (as described above in Section 2.1). Moving beyond just RV rms, John et al. (2023) also demonstrate that, for HD 144579, planets with periods between 1 and 2000 days have a mean detectability limit of just 0.62 m s^{-1} with RVs from the HARPS-N pipeline, and a mean detectability of 0.54 m s^{-1} for RVs that have been post-processed with *SCALPELS* (Collier Cameron et al. 2021).

The HARPS-N Low-Cost Solar Telescope (LCST; Dumusque et al. 2015; Phillips et al. 2016) is a 75 mm achromatic lens telescope with a fiber feed into HARPS-N. It has been observing the Sun as a star since first light in 2015 July. The HARPS-N LCST starts observing the Sun when it rises above an altitude of 20°. Calibrations are taken at the end of the previous observing night. There is a hard stop to observations when the Sun is below an altitude of 20°, but in practice observations are typically stopped every day between 14:00 and 16:00 UTC (15:00–17:00 local time) in preparation for nighttime observations. Data are calibrated using simultaneous observations of a Fabry–Perot etalon. Exposure times are fixed at 300 s (5 minutes) in an effort to bin over RV variations from solar oscillations (e.g., Chaplin et al. 2019). On a clear day, HARPS-N reaches an approximate analytical error per observation of 25 cm s^{-1}.

The pipeline for the HARPS-N LCST data is described in Dumusque et al. (2021). As with HARPS, the same data quality factor and atmospheric extinction coefficient are calculated in order to ensure good-quality data with HARPS-N (Collier Cameron et al. 2019). For HARPS-N, an additional quality cut is made using the ratio between the max and mean counts of the exposure meter for each observation. Observations with steady illumination of the fiber will return a ratio closer to one. Observations with a ratio greater than 3σ away from the mean are cut.

2.3. EXPRES

The EXtreme PREcision Spectrograph (EXPRES) is a stabilized, fiber-fed, optical (390–780 nm) échelle spectrograph (Jurgenson et al. 2016). It was fully commissioned at the 4.3 m Lowell Discovery Telescope (LDT; Levine et al. 2012) near Flagstaff, AZ in 2019 January. EXPRES has a median resolution of $R \sim 137,000$ and an instrument calibration precision of 4–7 cm s^{-1}, as measured by the rms of the perceived shifts of 0.5–1 hr of consecutive calibration exposures (Blackman et al. 2020; Zhao et al. 2022). For the quietest stars, EXPRES returns an on-sky precision of ~60 cm s^{-1} (Zhao et al. 2022).

The Lowell Observatory Solar Telescope (LOST) is a 75 mm lens telescope with a fiber feed into EXPRES that has been observing the Sun as a star since first light in late 2020 (J. Llama et al. 2023, in preparation). Calibration images for reducing solar observations are initiated when the Sun rises above 10°. Calibration images take about one hour to complete, and solar observations start immediately afterward; at this point, the Sun tends to be at an airmass of 2.6–3.0 depending on the time of year. Observations stop every day at 00:00 UTC (17:00 local time). Exposures are terminated at an S/N of 500, which on clear days corresponds to an average exposure time of ~180 s (3 minutes). At an S/N of 500, the analytic RV uncertainty per observation is about 35 cm s^{-1}.

Solar observations from EXPRES are extracted using largely the same pipeline for nighttime observations (Petersburg et al. 2020). Exposures are calibrated using lines generated from a Menlo Systems laser frequency comb (LFC; Wilken et al. 2012; Molaro et al. 2013; Probst et al. 2014, 2020; Milaković et al. 2020), which ranges from ~490 to 730 nm, and a hierarchical, nonparametric wavelength solution (Zhao et al. 2021). RVs are calculated using only the wavelength range for which there is LFC data. Dark and flat exposures are taken at the beginning of every day. Exposures are extracted using flat-

relative optimal extraction (Zechmeister et al. 2014). ThAr and LFC observations are taken every 30–45 minutes throughout the day.

For each observation, a quality factor is calculated using the standard deviation of the exposure meter counts, σ_{EM}, and the median of the exposure meter counts, Md_{EM}, as described further in J. Llama (2023, in preparation). The quality factor punishes exposures with large σ_{EM} values and large σ_{EM}/Md_{EM} ratios. RVs for each exposure were also regressed against time from solar noon to account for a daily offset thought to be due to the telescope flipping at solar noon (see J. Llama 2023, in preparation, for more discussion of this issue). Other than this regression, no correction is made for the differential atmospheric extinction of the Sun's resolved disk as it sets.

2.4. NEID

The NN-explore Exoplanet Investigations with Doppler spectroscopy (NEID) is a stabilized, fiber-fed, optical-to-NIR spectrograph with a full wavelength range of 380–1046 nm spectrograph (Halverson et al. 2016; Schwab et al. 2016). It was commissioned at the WIYN 3.5 m telescope[36] at Kitt Peak National Observatory near Tuscon, Arizona during 2020 and early 2021. NEID has a median resolution of $R \sim 120,000$, a total instrumental error budget of $27\,\mathrm{cm\,s^{-1}}$ (Halverson et al. 2016), and an on-sky precision of $41\,\mathrm{cm\,s^{-1}}$ for solar observations (Lin et al. 2022).

The NEID Solar Feed is a 75 mm lens with a fiber feed into NEID that has been observing the Sun as a star since 2020 December (Lin et al. 2022). NEID starts solar observations every day at 16:31 UTC (9:31 local time) and continues tracking the Sun until 22:30 UTC (15:30 local time) so as not to interfere with the calibration sequence for nighttime observations. A set of standard calibrations is captured both preceding and following the solar observations. Exposures are 55 s. This results in a typical analytic RV uncertainty per observation of $\sim 23\,\mathrm{cm\,s^{-1}}$ across all solar observations (we note that the $\sim 16\,\mathrm{cm\,s^{-1}}$ average analytic error cited in Table 1 is different because it was calculated using just the month of data being analyzed in this work).

Solar observations from NEID are extracted using the same pipeline as is used for nighttime observations; this pipeline is described at https://neid.ipac.caltech.edu/docs/NEID-DRP/.[37] For the RVs used here, the ESPRESSO G2 mask is used to calculate RVs, meaning only lines within the wavelength range of the mask (380–785 nm) are used to derived the NEID DRP RVs. Exposures are calibrated using a combination of simultaneous observations of a Fabry–Perot etalon, which ranges from 440 to 930 nm, and nonsimultaneous observations of a Menlo Systems LFC, which ranges from ~ 500 to 930 nm.

The NEID calibration protocols were developed for nighttime observations. In particular, a liquid nitrogen dewar is filled each morning, shortly after starting solar observations. NEID solar observations start at 16:31 UT; the liquid nitrogen refill is triggered at 17 UT. The dewar is filled in the morning so that any thermal transients, which can cause the instrument to experience large rapid drifts of $\sim 2\,\mathrm{m\,s^{-1}}$, will be settled during

nighttime observations. In contrast, the timing means the instrument drift has not settled during the initial part of the solar observations. The drift is tracked by the simultaneous calibration and is mostly—but not entirely—removed by the DRP. Many of the solar observations taken at the beginning of a day are therefore affected by nonlinearities in the wavelength calibration that render the resulting velocities less precise. Because the drift model was developed and tested for nighttime observations, it is likely that a future iteration of the pipeline could improve the wavelength calibration of the daytime observations.

To determine good-quality data (i.e., data for which the majority of the light in the exposure came from the full, integrated disk of the Sun), additional quality cuts were made on the NEID data analyzed here. These quality cuts made use of the time of day of the observation, data from a pyrheliometer, and the exposure meter counts. Early observations are cut to avoid the nonlinear behavior that follows the dewar filling; late observations are cut to avoid outliers seen on some days, likely due to exposures continuing after the solar feed stops tracking and/or is obstructed. In addition to the exposure meter, NEID uses a pyrheliometer to provide an independent measure of the Sun's irradiance. The standard deviation and mean of both the pyrheliometer measurements and exposure meter counts are used to assure observations were taken during a time of steady and significant solar flux. The exposure meter flux is then compared to the pyrheliometer flux to ensure they agree, indicating good pointing during the exposure.

3. Methods

DRP observations are binned to a common set of time stamps, to allow for a direct comparison between the solar data collected by each instrument. Using the binned RV time series, we can determine the agreement between instruments both within and across a day (Section 3.2). We additionally measure day-to-day offsets for each instrument (Section 3.3).

For both the binned RVs and the day-to-day offsets, we assume the provided analytic errors give 1σ Gaussian errors for each data point, and we use a Monte Carlo approach to determine empirical errors. The given analytic errors are used to generate 1000 random samples of the observed time series. For each observation, a new random RV is drawn from a normal distribution with a mean of the original RV value and a sigma of the given analytic error for that observation. Each of the 1000 randomly generated time series is then binned using the same method as described below. The scatter in the resulting binned point across all simulations is used as the error for the binned point.

The expected amplitude of various sources of error is given in Table 2. For each observation, we have analytic errors from the corresponding instrument's DRP. The average analytic error as reported by each instrument's DRP is given in Table 2 under "Analytic." The average error of the binned RVs for each instrument, estimated as described above, is given in Table 2 under "Binned."

3.1. Binning

The RVs collected by each instrument are interpolated onto shared time stamps to allow for a direct comparison. The Sun is expected to exhibit p-mode oscillations with a ν_{max} of

---

[36] The WIYN Observatory is a joint facility of the NSF's National Optical-Infrared Astronomy Research Laboratory, Indiana University, the University of Wisconsin–Madison, Pennsylvania State University, and Purdue University.

[37] All data from the NEID solar feed are made available in a timely manner at https://neid.ipac.caltech.edu/search_solar.php.

Table 2
Average Error from Different Sources

Instrument	Analytic[a]	Binned[b]	Compared with	No. Points[c]	Binning Rms[d]	Residual Rms[e]	Remaining Scatter
HARPS-N	29	21					
			HARPS	161	11	36	17
			EXPRES	91	26	47	33
			NEID	10	10	20	…
HARPS	30	9					
			EXPRES	306	25	43	26
			NEID	229	14	26	17
EXPRES	19	12					
			NEID	508	26	39	16
NEID	16	6					

Notes. All values are given in units of cm s^{-1}.
[a] Average analytic error over the month of shared data.
[b] Average empirical error of binned RVs.
[c] Total number of overlapping points.
[d] Average scatter expected from imperfect binning over p-modes.
[e] Scatter from the residuals of binned RVs.

5.4 minutes (3.1 mHz), which introduces RV variations on that timescale (Kjeldsen et al. 2008). The different time stamps and observing cadence of the four instruments will therefore be sampling this high-frequency signal at different points. The variation in exposure times used by the different instruments will also average out this signal to different degrees.

We ran tests on simulated p-mode oscillation time series, to compare different methods of binning and to estimate the expected error from the binning process. Oscillations are simulated as a series of damped driven simple harmonic oscillators (SHOs). We include the contributions from 32 SHOs corresponding to $\ell = 0$, 1 modes. More SHOs and higher-order modes were found not to contribute significantly to the power spectrum. The characteristic frequencies and amplitudes of the SHOs were set to approximate the observed solar p-mode power spectrum. Specifically, we use a central peak frequency of $\nu_{max} = 3100\ \mu$Hz, a peak separation of $\Delta\nu = 1.3 \times 10^{-4}$ Hz, and a Gaussian envelope width of $331\ \mu$Hz, in accordance with the values adopted by Chaplin et al. (2019). The damping timescale was set to two days, a typical value for solar-like stars (Chaplin et al. 2009). The amplitude of the stochastic driving force was set to produce a coherent RV amplitude of $A_{max} = 0.19$ cm s^{-1}, again in accordance with Chaplin et al. (2019). At each time stamp in a continuous grid with one-second separations, we solve for the position and velocity of each damped SHO component, apply a stochastic driving force in the form of a randomized velocity perturbation, and sum the velocity contributions from each oscillator (or "mode") to yield a simulated asteroseismic RV.

These simulated p-modes are then turned into realistic, observed time series that mirror the specifics of the data sets from each instrument. Observations are derived from the simulated p-modes by first interpolating between the start and end time of an observation onto a finer, 0.1 s grid. This finely sampled signal is averaged to produce an "observed" RV from the simulated p-mode time series. The time stamp and start/end time of each simulated "observed" RV exactly match the properties of the real observations from each instrument. In this way, the simulated p-modes are converted into observations that match the properties of the real observations being analyzed.
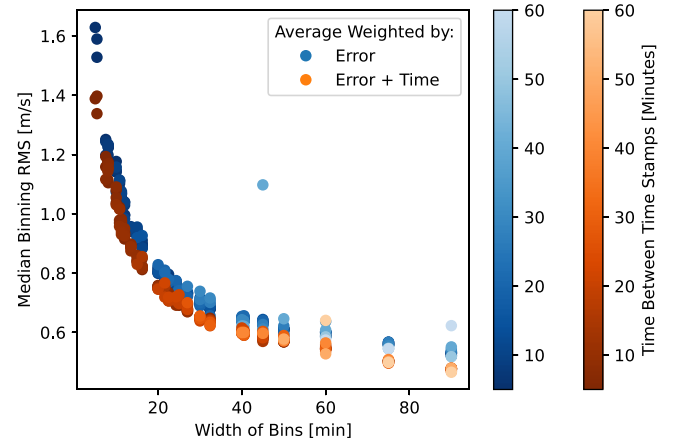

Figure 5. Resulting Binning rms from different binning methods as a function of bin width. The color of the point indicates whether the RVs in a bin were weighted by analytic error before being averaged (blue) or weighted by error and time from the center of the bin (orange). The saturation of the color indicates the time between each bin.

We generated 100 realizations of "observed" time series from simulated p-mode oscillations to test a variety of binning methods. Binning methods varied in time between each shared time stamp (between 5 and 60 minutes) as well as the width of each bin (between 2.5 and 90 minutes), where the bin width was constrained to be greater or equal to the time between each time stamp. The data in each bin were combined via a weighted average where the weights were determined using either (1) just each observation's given analytic error or (2) each observation's error as well as the separation in time between the observation and the center of the bin.

Errors for each binned data point were derived using a Monte Carlo (MC) method. For 1000 simulations, the true DRP observed RVs were perturbed by their given 1σ analytic errors (the average analytic error is given in Table 1 and repeated in Table 2 under the "Analytic" column). These perturbed points were then binned. The resultant scatter in the binned points across all 1000 simulations was taken to be the MC-derived error for the binned points. This error is give in Table 2 under the "Binned" column.
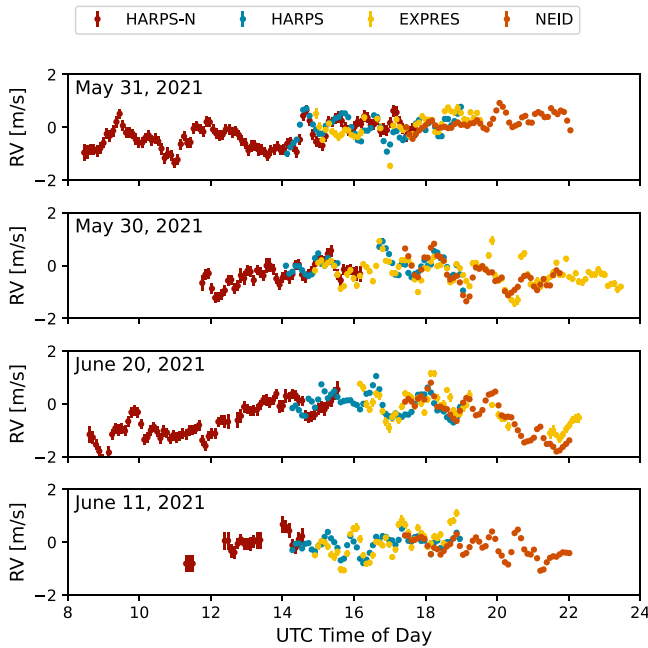
Figure 6. Binned RVs as a function of UTC time of day. We show here the four days for which there were the greatest number of overlapping points between all instruments. Error bars for each binned RV are derived empirically, as described in Section 3. A different RV offset is applied for each instrument for each day; the offset is calculated as the median of the binned RVs for that instrument taken on the day in question. Features that are traced by more than one instrument are most likely solar in nature rather than due to instrument-specific systematics.

Because the point of the binning is to allow for a direct comparison of the solar data gathered by each instrument, the binning methods were evaluated on the scatter returned by the resulting binned RVs when a pair of instruments were differenced. We take the spread in residuals from these comparisons of simulated observations as the expected error due to the binning process alone; these values are given in Table 2 under "Binning rms." The median Binning rms across the simulated 100 p-mode time series from the different binning methods is shown in Figure 5 as a function of bin width and colored by bin spacing and weighting method.

The resultant "Binning rms" decreases with the width of each bin following a power law. This is expected because wider bins allow for more data to be averaged into a shared point. The resultant rms values are consistently $\sim 5\,\mathrm{cm\,s^{-1}}$ lower when RVs are weighted by RV error and time of observation (as opposed to just RV error). The shape of the power law is dominated by the width of the RV bins and is unaffected by the spacing of the time bins. Fitting the relation to a power law reveals no significant deviations that would suggest a given binning method is doing better than what is expected when more data is added.

Because there was no clear evidence of a given binning method returning smaller scatter than expected, we chose a binning method based on what is understood of the Sun and the data used in this analysis. We establish time stamps that are spaced 5.4 minutes apart, the ν_{max} for the Sun. For each time stamp, all observations within a $(5.4 \times 3=)16.2$ minutes window are combined via a weighted average. RVs are weighted by both the inverse variance of the observation and a linear weighting based on the time from the center of the bin. For example, a point falling exactly at the center of the bin

would get a relative weighting of 1, and a point at the edge of the bin (i.e., $16.2/2 = 8.1$ minutes away) would get a relative weighting of zero. The longer, 16.2 minute window ensures that there are enough observations in a bin to produce a reasonable average, even for observations with the longest exposure times (5+ minutes). Some observations will be included in multiple bins, resulting in a smooth binned behavior as expected. Binned RVs are shown in Figure 6 for the days on which there is the greatest amount of overlap between the four instruments.

3.2. Direct Comparison of Binned RVs

Binning observations from each instrument to shared time stamps allows us to directly compare binned RVs between instruments. Before comparison, an RV offset, calculated as the median of all binned RVs, is removed from each time series. Figure 7 shows the spread in the residuals when the binned RVs from different pairs of instruments are differenced. Each of the six columns corresponds to a different pair of instruments, as given at the top of each column, where the total number of overlapping points between the two instruments is written at the top of the plot.

This spread in residuals, as written on the top of each subplot, is also given in the table of errors (Table 2) under "Residual rms." The distribution of residuals from the simulated p-mode oscillation is overplotted in black. In other words, the expected spread from just the error introduced from binning is shown by the black outline histograms.

The bottom row of the figure shows the binned RVs from each instrument plotted directly against the binned RVs from each of the other instruments. The Pearson correlation coefficient, ρ, between the two binned RV time series is given in the top left of each subplot.

The last column of Table 2 gives the magnitude of the remaining unexplained scatter. As described, we have quantified an error due to (1) photon noise, as given by the analytic error for each data point, and (2) the binning process, as determined using simulated p-mode oscillation time series. Assuming these two noise sources are independent, Gaussian, and make for a complete accounting of the noise present in the comparison, we would expect the residual rms to be the analytic error and the binning rms added in quadrature. In actuality, we find excess noise in the residual rms, which is given in the "Remaining Scatter" column of Table 2 (i.e., this is the amount of error that still has to be added in quadrature, in addition to the analytic and binning error, to obtain the residual rms). For the comparison of HARPS-N with NEID, no remaining scatter value is given, because there are too few data points (only 11).

Figure 8 shows pairwise residuals for each instrument phase-folded to the UTC time of day. For each of the four instruments, the residuals with respect to all three of the other instruments are shown. Phase-folding the residuals to time of day is useful to emphasize nonastrophysical effects. We would not expect the Sun to exhibit any coherent variation with respect to time of day on Earth. Therefore, any coherent structures seen as a function of time of day is likely due to Earth-bound observing conditions or daily data processing effects.

Features seen in the residuals with respect to more than one instrument are of particular interest. Having the same coherent variation appear in more than one comparison provides
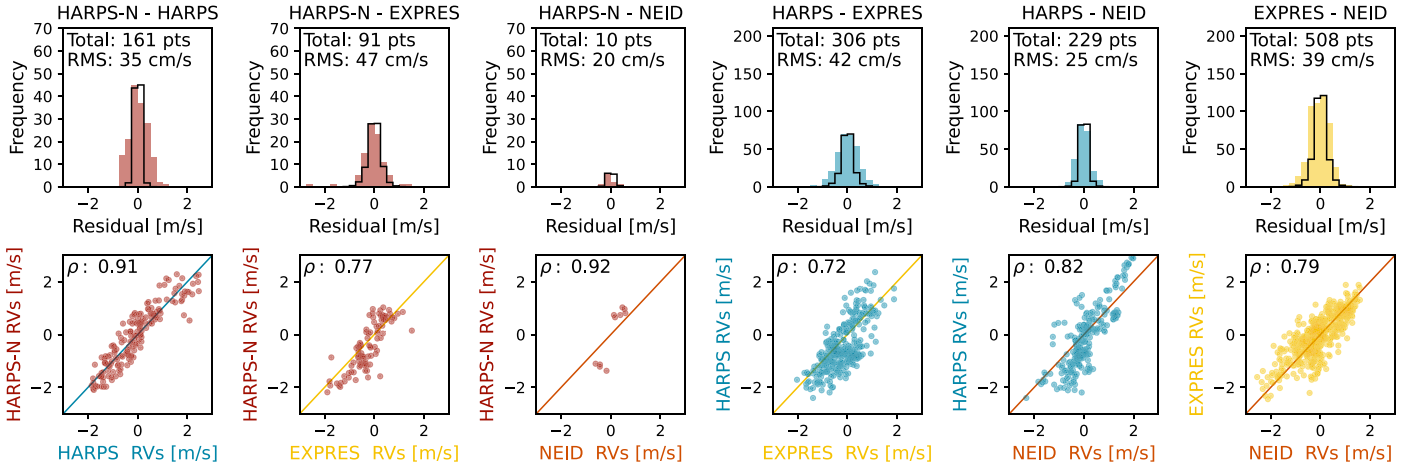
Figure 7. Top: histograms of the residuals between one instrument's binned RVs against another instrument's binned RVs. The title of each subplot details the two instruments being compared (where the binned RVs of the second instrument listed are subtracted from the binned RVs of the first instrument). The total number of overlapping binned RVs across the entire month of shared data is given at the top of each subplot, followed by the rms of the residuals. Bottom: Comparison of one instrument's binned RVs (y-axis) against the binned RVs of the other three instruments (x-axis). A point is shown for each binned time stamp at which both instruments have a binned RV. The corresponding Pearson correlation coefficient, ρ, for each pairwise comparison is given in the top left corner. Values of ρ closer to 1 indicate more significantly positively correlated data.
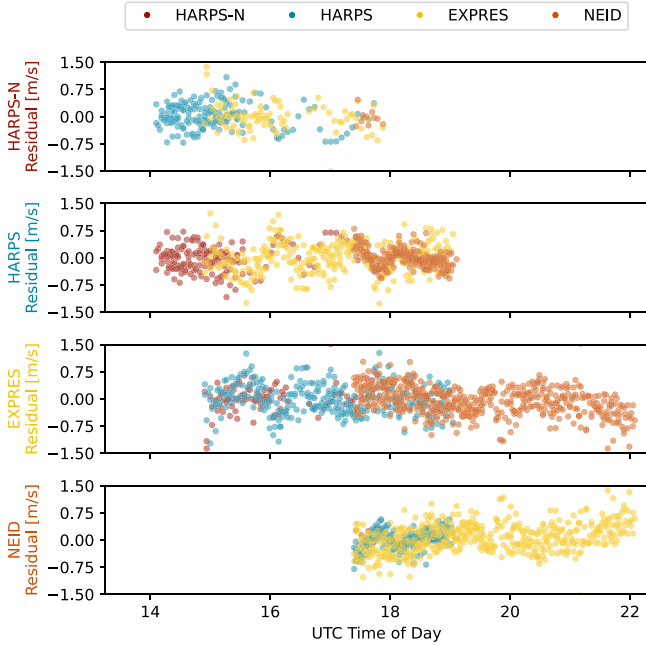
Figure 8. Residuals between binned RVs across all 29 days phase-folded to UTC time of day. Each subplot shows the residuals between binned RVs of the instrument specified in the y-axis label and the other three instruments. A point is shown for each binned time stamp at which both instruments have a binned RV. Of particular interest are features traced by the residuals from multiple instruments, which lends confidence to the feature and helps diagnose its origin.

independent confirmation of the variation. Additionally, with pairwise comparisons, having the same features appear in multiple comparisons is needed to pinpoint the origin of the feature. As an example, if a comparison between instruments A and B reveals a coherent variation, it is unclear if it is instrument A, instrument B, or a combination of the two causing that coherent variation, in the absence of ground truth. However, if a comparison between instruments A and C exhibits the same coherent variation, then we are able to conclude that instrument A is the origin of the variation.

3.3. Day-to-day Offsets

We measure day-to-day offsets (i.e., a single RV offset for each day of solar data) for each instrument relative to the other instruments. We know there exists measured RV variations due to changes on the solar surface. Ideally, it would be possible to construct a model of these variations. The RVs from each instrument for each day could then be compared to this model, giving an offset for each day. With the current data, it was not possible to get a good fit for such a model. Because there is a minimum eight-hour gap in solar data every day (between when the Sun set for the westmost instrument and when the Sun rose again for the eastmost instrument), there is little constraint on the model from one day to the next.

Rather than prescribe a potentially flawed model of the expected RV variations, we compare each instrument to the RVs measured by the other instruments. Because all instruments are observing the same solar surface variations, it is reasonable to use the stellar signals measured by another instrument in place of a model of the expected RV variations. This, of course, ignores the inherent error of each observation from photon noise, instrument systematics, and varied time stamps/exposure times.

To measure the offset between two instruments over a day, we find the difference between the median of the binned RVs, where only the binned RVs where the two instruments overlap are used to calculate this median. The resultant day-to-day offsets are shown for each instrument in Figure 9. Points are only shown for days where the two instruments had overlapping binned RVs. We note, however, that the number of overlapping points differs between instruments and between days, meaning the offsets are not directly comparable. Offsets on days for which there exist offsets calculated relative to two or more instruments are shown as filled-in points. Of particular interest are days on which multiple offset calculations are consistent with one another. Measurements from different comparisons that agree lends confidence to the measured offset for the instrument in common between the comparisons.

The error bars in Figure 9 are derived using the simulated p-mode time series described above in Section 3.1. The resultant observed time series derived from the simulated
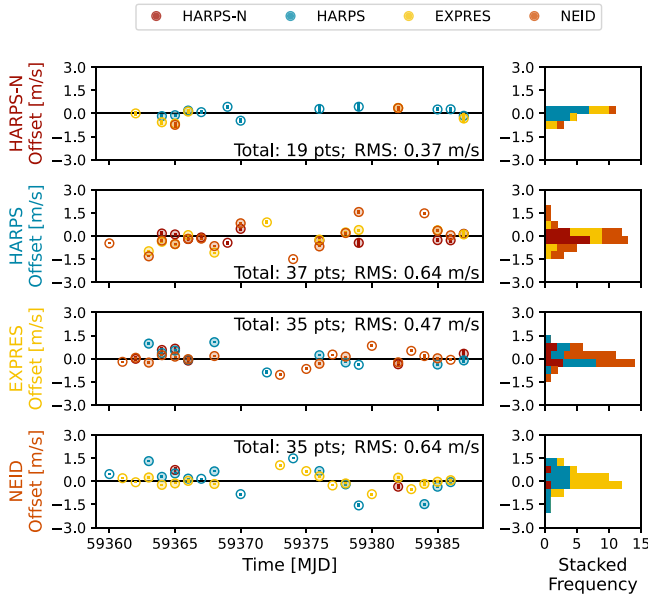
Figure 9. Left: day-to-day offsets relative to other instruments. Each subplot shows the day-to-day offsets for the instrument specified in the *y*-axis label as calculated relative to the other three instruments. A point is shown for each day in which both instruments have at least one overlapping binned RV. Days for which there exist offsets measured against two or more instruments are shown as filled-in circular points; days for which there only exists one comparison are shown as outlines. It should be noted that the expected error due to binning for each point is shown as an error bar, but all error bars lie interior to each point, suggesting that binning error is not the dominant source of uncertainty. The total number of days for which an offset was calculated across all three comparisons as well as the rms of all offsets are written on the right of each subplot. Right: stacked, horizontal histograms of the relative day-to-day offsets calculated for each instrument. The resulting distribution of comparisons against each of the three instruments is stacked on top of each other. In other words, the height of a multicolored bar is the frequency of that bin's values across the comparison with multiple instruments.

p-modes are used to calculate day-to-day offsets for each instrument as described. The spread in the returned offset across all 1000 simulations is shown as the error in Figure 9. It should be noted that this error only reflects the expected error from binning p-modes into discrete observations and then to shared time stamps. It does not capture errors that would, for example, arise from instrument systematics, varied levels of differential extinction, or any other source of uncertainty that is not related to binning.

4. Results

We are able to probe instrument-to-instrument precision within a day, across a day, and across the 29 days of shared data. Precision within a day measures the degree to which each instrument returns the same RV variation in response to solar surface changes. Should the agreement between instruments vary across a day, this may imply day-long instrument drifts or imperfect correction for daily effects, such as atmospheric extinction or instrument drift. Instrument precision and accuracy from day to day hints at the long-term stability of each instrument—and consequently an instrument's ability to recover long-period signals.

4.1. Precision within a Day

RVs from multiple instruments often trace the same features. It is easy to see this in Figure 6 where the binned RVs from

each instrument are shown for the four days on which there is the greatest amount of overlap. Though the binning cannot completely mitigate the differences in RV that arise simply from different time stamps, we see here that the binned RVs from each instrument do seem to largely agree on the resultant measured RV variation from solar surface changes.

While the bin widths average over p-mode oscillations (Chaplin et al. 2019), there remains short-timescale oscillations in the binned RVs, likely due to granulation. Given that the same oscillatory features are traced by multiple instruments, they are likely astrophysical in nature. Similar variations have been found in HARPS astroseismology measurements and in an analysis of three years of HARPS-N solar data (Dumusque et al. 2011b; Collier Cameron et al. 2019). These analyses concluded that such short-timescale variations are likely due to granulation phenomena.

Histograms of the residuals between the binned RVs of pairs of instruments (top row of Figure 7) show the agreement between all overlapping binned RVs. The expected error from comparing binned RVs between instruments is given in Table 2 in the "Binning rms" column. Comparisons to the EXPRES binned RVs give rise to the largest expected errors. Of the four instruments, EXPRES is the only instrument that makes use of adaptive exposure times, exposing to a set S/N rather than a set time. Because we expect the error from binning to scale directly with exposure time, having changing exposure times may change the expected binning error throughout the EXPRES time series, leading to these larger "Binning rms" values.

After taking into account the expected error from generating binned RVs to the shared time steps, we find that the pairs of instruments agree with each other to within 15–30 cm s^{-1} (see the "Remaining Scatter" column of Table 2). This shows that, even using the default DRP derived RVs, the instruments are in very good agreement with one another. The bottom row of Figure 7 plots the binned RVs from the two instruments against each other. The RVs are linearly correlated and have significant correlation coefficients in all cases, as expected.

4.2. Precision across a Day

To investigate variations in instrument precision across a day, we show the residuals for each instrument phase-folded to UTC time of day in Figure 8. Any coherent structure as a function of time of day is most likely a cause of necessary daily corrections—e.g., to account for things such as instrument drift, or Sun-specific effects such as the changing atmospheric extinction as the Sun rises and sets, etc.—rather than true variations on the solar surface.

There appear to be oscillations in the HARPS residuals (see the second subplot from the top of Figure 8). These oscillations have a semimajor amplitude of 85 cm s^{-1} and a period of 30–50 minutes. The oscillations show most clearly in the residuals as compared to EXPRES data (yellow points), for which there is the greatest amount of overlap with the HARPS data, but the same oscillations also appear in the residuals to the NEID RVs (orange) and possibly the HARPS-N residuals (red) as well. We can therefore attribute this behavior to the HARPS measurements. This feature is only seen during a short period of the HARPS solar observations, including the period analyzed in this paper. It is likely due to an unusual amount of dust on the telescope transparent dome that induced differential extinction. This effect will be investigate further in a forthcoming paper.

There appears to be a slope in the residuals between the EXPRES and NEID RVs when the Sun would be setting for these two instruments. After UTC time of day $\sim$21, NEID RVs are increasingly higher, or more redshifted, than EXPRES RVs as the Sun sets. The slope of this trend is approximately 73 cm s^{-1} per hour as fit using data taken after a UTC time of 21, where the slope is the steepest. This may indicate a need to account for high-airmass effects or differential extinction. Though neither instrument explicitly accounts for differential extinction, the solar noon correction applied to the EXPRES observations (as described in Section 2.3) likely mitigates differential extinction to some degree. HARPS and HARPS-N do explicitly account for differential extinction (as described in Sections 2.1 and 2.2) and do not show a similar linear trend. This trend can therefore in principle be attributed to whether or not the effects of high airmass and differential extinction are corrected for.

4.3. Stability across Days

RVs from multiple instruments often trace the same long-term features as well as daily ones. Figure 1 plots the binned RVs from all instruments in the bottommost subplot. The RVs from each instrument generally follow the same trend throughout the month.

Figure 9 shows the derived day-to-day offsets for each instrument relative to the other instruments. We chose to measure offsets relative to the other instruments in order to avoid introducing any potential errors from assigning a flawed model of the expected measured RV variations. This pairwise comparison, however, comes with its own caveats.

The measured daily offsets are derived using variable amounts of data. An offset is shown for any day on which there exist overlapping RVs for two instruments. The number of binned RVs used to calculate the offset therefore varies from a single pair of overlapping binned RVs to nearly 60 overlapping points in some cases. The error bars shown for each point, which are on the order of 5–25 cm s^{-1}, capture only the expected errors from binning and so do not reflect the different amounts of information going into each of the points. Because of the varied amounts of overlapping data from day to day, it is difficult to directly compare the measured offsets with one another, as they differ widely in information content.

There do exist some days for which the day-to-day offset for an instrument is reflected in the comparison against more than one instrument. Because all offsets are relative and pairwise, it is hard to tell from a single comparison what the source of the offset is. If comparisons against two different instruments return the same offset, however, that lends credibility to the returned offset. We therefore emphasize days for which an instrument has a calculated offset relative to two or more instruments by showing these days with filled-in circular markers in Figure 9.

Figure 9 shows stacked histograms to visualize the distribution of the day-to-day offsets calculated for each instrument. The histograms are stacked to show the overall distribution of all calculated offsets. The colors of the bars correspond to the instruments used to calculate each day-to-day offset. This makes it easier to determine if any skewness in the distribution is due to one instrument in particular. For example, the orange-colored bars in the third plot represent day-to-day offsets for EXPRES as compared to NEID data. The orange bars are primarily responsible for the stacked histogram

appearing to be skewed toward positive offsets. This matches our finding from Figure 8 that residuals between EXPRES and NEID binned RVs are skewed positive as the Sun sets.

We note that that this way of calculating the day-to-day offset is particularly sensitive to differential extinction when comparing instruments in different time zones. Not only are the effects of differential extinction greatest when the Sun is rising or setting, the effect is also opposite between the Sun rising and the Sun setting. For instruments with a large separation in time zones, many of the overlapping data points will be when the Sun is setting in one location and rising in the other. The spread in the daily offsets between HARPS-N as compared to EXPRES or NEID (shown by the yellow and orange bars in the second row of the right column of Figure 9) is larger than the spread in the daily offsets between EXPRES and NEID. This could be because EXPRES and NEID are geographically close and therefore view the Sun concurrently at a similar airmass, while HARPS-N is further to the east and south.

We find that the scatter in day-to-day offset for each instrument is on the order of 50–60 cm s^{-1}. This is likely artificially inflated, given the shortcomings of using relative offsets as described above. More accurate estimates can be achieved with data that more completely span the 24 hr in a day and more days than the 29 used in this analysis.

5. Discussion

Combining solar data from different instruments makes for a powerful test bed on which to understand methods for mitigating stellar signals. With solar data, it is possible to confidently remove measured velocity shifts due to Solar System bodies. Because these solar feeds collect sunlight in a way that closely emulates light collection of other stars, the resultant data set captures real stellar surface variations in the same manner as would be observed on other stars. High-S/N and high-cadence observations are uniquely easy to obtain, and it is simple to degrade these observations to approximate the lower-S/N and sparser observations of nighttime targets. By combining observations from multiple instruments, we are also able to probe systematics that are unique to individual instruments.

There are also limitations of observing the Sun. As a resolved source, differential extinction as the Sun rises and sets causes spurious RV shifts at high airmass, something that does not happen to other stars, which appear as point sources (Davies et al. 2014; Collier Cameron et al. 2019). Because the Earth orbits the Sun, the barycentric corrections for solar data are of a much smaller magnitude than when observing other stars. Telluric lines therefore shift with respect to stellar lines to a lesser degree, making it far more difficult for methods that depend on a large difference in stellar and barycentric shift to be tested on solar data (e.g., Bedell et al. 2019; Cretignier et al. 2021). Solar light is often injected into the fiber leading to the spectrograph via a different optical path than nighttime observations, which are taken with the main telescope. Analyses using solar data are therefore insensitive to possible instrument systematics unique to nighttime observing, such as the guiding and focusing of the main telescope. Last, the Sun is just one example of a star, and how its surface variations map to the surfaces of other stars with different spectral types and stellar properties has not been well characterized.

5.1. Standardizing RV Derivation

Here, we use the default pipeline RVs from each instrument and find that the different instruments agree very well with each other. The binned RVs from each instrument are strongly correlated with other instruments and can be visually seen to trace similar patterns. When taking into account the given analytic error and the error from binning, we find that a direct comparison between RVs from different instruments shows that they agree to within 15–30 cm s^{-1}.

It might be possible to get even better agreement between instruments by standardizing how the RVs are derived. HARPS, HARPS-N, and NEID all produce RVs using the cross-correlation function (CCF), where a mask is shifted to find the best match with the cores of a given list of spectral lines. The EXPRES RVs are derived using a forward-modeling framework that uses all spectral lines in the range where there is LFC light ($\sim$490–730 nm). The contribution of each line to the final RV is based on an empirical measure of the RV content of the line.

We expect different lines to respond differently to stellar signals. Therefore, we would expect RVs to agree the best when the same stellar lines are used to derive them. All CCF RVs used in this analysis were derived using the ESPRESSO G2 line list, and so they should have incorporated many of the same lines across a similar wavelength range. We also find that the blaze function of each instrument peaks at similar locations for all instruments (see Figure 3), meaning the relative weighting of a spectral line across an order will be similar for all instruments.

Despite having similar line lists and throughput across an échelle order, the pipeline RVs still differ in many ways. EXPRES RVs are derived using a forward-modeling method that only incorporates wavelengths for which there is LFC coverage ($\sim$490–730 nm). Even for HARPS, HARPS-N, and NEID, which all use the CCF method and the ESPRESSO G2 mask (with lines between $\sim$380 and 785 nm) for deriving RVs, the CCF is calculated by different pipelines.

Different pipelines also differ in the amount of information per order and the way the RV information content from each order is combined. The size of the detector differs between instruments, changing the wavelength range of each individual échelle order—and consequently the number of spectral lines that appear in more than one order. The overall throughput as a function of wavelength also varies by instrument. The relative signal of a line in the blue as compared to a redder line therefore changes with instrument. The same CCF mask is used by HARPS, HARPS-N, and NEID, to derive a CCF for each order, but each order's CCF/RV is weighted and combined into a global RV for an observation in a different way by the different pipelines.

The different pipelines avoid telluric contamination in different ways. EXPRES uses a telluric model derived by SELENITE (a SELf-calibrating, Empirical, Light-Weight liNear regressIon TElluric model; Leet et al. 2019) to identify and mask out telluric lines. Though HARPS, HARPS-N, and NEID all use the ESPRESSO G2 CCF mask, the different pipelines have different cutoffs for avoiding shifting line masks over telluric lines.

It may therefore be instructive to run the same exact pipeline on solar data collected by different instruments in order to eliminate pipeline differences and isolate stellar and instrumental effects. Because we find that the pipeline RVs are themselves in fairly good agreement, we leave investigating potential improvements by standardizing RV derivation to future work.

5.2. Benchmarking Instruments against Instruments

Concurrent observations of the same source from four different instruments allowed us to benchmark instruments against each other and gain a deeper understanding of the observations. Concurrent observations removed the need for constructing potentially flawed models to compare against. However, the different time stamps and exposure times of observations introduced errors in how high-frequency signals were sampled within each time series. This would not be an issue for longer-period signals (e.g., on the timescale of planet periods or stellar rotation rates). Instrument-to-instrument comparisons were also only possible when concurrent observations fell very close to each other; data with observational gaps and/or telescopes geographically far apart reduce the amount of overlapping data that can be used to inform a rigorous comparison.

The differential extinction of integrated sunlight as the Sun rises and sets further complicates comparisons between instruments. For geographically distant instruments, much of the overlapping data will be taken as the Sun is setting in one location and rising in another. Comparisons between instruments will then mostly use observations taken at high airmass and subject to the effects of differential extinction in opposite ways (i.e., the difference between the effect as the Sun rises versus sets). This is likely to artificially inflate the scatter in the residuals and calculated day-to-day offsets.

We find the spread in the day-to-day offset for each instrument to be on the order of 50–60 cm s^{-1}, which is much larger than the intraday residual scatter of 15–30 cm s^{-1}. It is possible that instrument systematics or data reduction artifacts truly cause the RVs from each instrument to change by up to 60 cm s^{-1} from day to day. However, it is more likely that the calculated day-to-day offsets are artificially inflated due to differential extinction, which can cause an RV difference of about 50 cm s^{-1} at an airmass of 2.2 (Collier Cameron et al. 2019), an effect that will compound if observations taken when the Sun is setting are compared to observations taken by a different telescope when the Sun is rising. Differences may also be mitigated through more standardized pipelines or observing strategies. It will be difficult to reach a significant conclusion about the day-to-day offsets for different instruments without more overlapping data that are taken with comparable time sampling.

Comparing solar telescopes in the same time zone allows for more confident comparisons. In this analysis, EXPRES and NEID are in the same time zone and share by far the greatest number of overlapping data points. It will soon be possible to perform a similar analysis with HARPS and HARPS-N data as solar telescopes for ESPRESSO (PoET; Leite et al. 2022)[38] and HARPS-3 (ABORAS; Farret Jentink et al. 2022) come online, covering the same time range as HARPS and HARPS-N, respectively.

Wider time coverage of solar data will also allow for a better estimate of day-to-day offsets. There exist global, ground-based networks of solar observatories, such as the Global Oscillation Network Group (Harvey et al. 1996) and the

---

[38] See http://poet.iastro.pt for more information.

Birmingham Solar Oscillations Network (BiSON; Davies et al. 2014; Hale et al. 2016). It is worth exploring if the single-line spectra produced by these networks can provide insight into the expected RV variations as measured by the solar telescopes investigated here. The space-based Solar Dynamics Observatory (Pesnell et al. 2012) also provides near-continuous images of the Sun at a variety of wavelengths as well as a measure of the Sun's magnetic activity, which could be used to inform the expected RV response from solar surface variation (e.g., Ervin et al. 2022). It is also possible, in principle, to collect solar data at night via reflected light from the Moon or asteroids to increase overlap between instruments.

This analysis is imminently repeatable and is guaranteed to only get more interesting. Over the next few years, several new solar feeds for precision spectrographs are expected to come online. PEPSI (Strassmeier et al. 2015) and GIANO-B (Claudi et al. 2018) also have solar feeds. Solar feeds for KPF (R. Rubenzahl 2023, in preparation), ESPRESSO (Leite et al. 2022), HARPS-3 (Farret Jentink et al. 2022), MAROON-X (Seifahrt et al. 2018), and NIRPS (Bouchy et al. 2017) are in the works. We note that shorter exposure times give rise to less error when binned to share time stamps, meaning shorter solar exposures will be easier to compare with observations from other instruments.

The observations themselves will also become more interesting as the Sun moves out of an activity minimum. The data used in this analysis were taken only 14 months after the latest cycle minimum, which occurred in 2019 December. Activity on the Sun is expected to peak late in the Fall of 2024 with a maximum sunspot number of 134 ± 8 (Upton & Hathaway 2023).

It is also, of course, possible to perform a similar analysis with observations of stars other than the Sun. Observations of other stars are less susceptible to differences in airmass and do not experience differential extinction as solar observations do. This would remove an error source in benchmarking instrument variations that was encountered in this analysis of solar data. Repeating such a comparison with more stars will also inform instrumental and stellar variation across different stellar properties.

From this analysis, it is clear that coordinating observation times will have a huge effect on the ability of different data sets to benchmark one another. Shorter observations are more easily binned to shared time stamps with less inherent error from binning over high-frequency variations. Given the presence of short-timescale variations, observations from different instruments should be taken as close in time to each other as possible. Therefore, targets that are bright and high in the sky will be easiest to observe in a way that is favorable for cross-instrument comparisons. As such coordination is sure to be logistically difficult, the consistency and high S/N of solar observations will continue to remain valuable even as cross-instrument analyses are pursued with other stars.

By combining solar data sets from different instruments, we were able to learn more about individual instrument systematics, confirm observations of solar signals, and produce a powerful data set for gauging the performance of different methods for mitigating stellar signals. We find that binned RVs from each instrument all agree to within 15–$30 \, \mathrm{cm \, s^{-1}}$ when analytic and binning errors are taken into account, even though the four instruments started operations almost 20 yr apart. This demonstrates the precision that the four instruments analyzed here are capable of, showing that small-amplitude signals can be confidently detected with existing instrumentation if stellar signals are successfully disentangled to the needed levels. This analysis showcases the potential of such a comparison, which is repeatable for a larger number of instruments, a longer time baseline of data, and/or even for stars other than the Sun.

ORCID iDs

Lily L. Zhao ⓘ https://orcid.org/0000-0002-3852-3590
Xavier Dumusque ⓘ https://orcid.org/0000-0002-9332-2011
Eric B. Ford ⓘ https://orcid.org/0000-0001-6545-639X
Joe Llama ⓘ https://orcid.org/0000-0003-4450-0368
Annelies Mortier ⓘ https://orcid.org/0000-0001-7254-4363
Megan Bedell ⓘ https://orcid.org/0000-0001-9907-7742
Khaled Al Moulla ⓘ https://orcid.org/0000-0002-3212-5778
Chad F. Bender ⓘ https://orcid.org/0000-0003-4384-7220
Cullen H. Blake ⓘ https://orcid.org/0000-0002-6096-1749
John M. Brewer ⓘ https://orcid.org/0000-0002-9873-1471
Andrew Collier Cameron ⓘ https://orcid.org/0000-0002-8863-7828
Rosario Cosentino ⓘ https://orcid.org/0000-0003-1784-1431
Pedro Figueira ⓘ https://orcid.org/0000-0001-8504-283X
Debra A. Fischer ⓘ https://orcid.org/0000-0003-2221-0861
Adriano Ghedina ⓘ https://orcid.org/0000-0003-4702-5152
Samuel Halverson ⓘ https://orcid.org/0000-0003-1312-9391
Shubham Kanodia ⓘ https://orcid.org/0000-0001-8401-4300
David W. Latham ⓘ https://orcid.org/0000-0001-9911-7388
Andrea S. J. Lin ⓘ https://orcid.org/0000-0002-9082-6337
Gaspare Lo Curto ⓘ https://orcid.org/0000-0002-1158-9354
Marcello Lodi ⓘ https://orcid.org/0000-0002-0651-4294
Sarah E. Logsdon ⓘ https://orcid.org/0000-0002-9632-9382
Christophe Lovis ⓘ https://orcid.org/0000-0001-7120-5837
Suvrath Mahadevan ⓘ https://orcid.org/0000-0001-9596-7983
Andrew Monson ⓘ https://orcid.org/0000-0002-0048-2586
Joe P. Ninan ⓘ https://orcid.org/0000-0001-8720-5612
Francesco Pepe ⓘ https://orcid.org/0000-0002-9815-773X
Rachael M. Roettenbacher ⓘ https://orcid.org/0000-0002-9288-3482
Arpita Roy ⓘ https://orcid.org/0000-0001-8127-5775
Nuno C. Santos ⓘ https://orcid.org/0000-0003-4422-2919
Christian Schwab ⓘ https://orcid.org/0000-0002-0091-7105

Guðmundur Stefánsson ⓘ https://orcid.org/0000-0001-7409-5688
Andrew E. Szymkowiak ⓘ https://orcid.org/0000-0002-4974-687X
Ryan C. Terrien ⓘ https://orcid.org/0000-0002-4788-8858
Stephane Udry ⓘ https://orcid.org/0000-0001-7576-6236
Sam A. Weiss ⓘ https://orcid.org/0000-0002-5870-8488
François Wildi ⓘ https://orcid.org/0000-0002-9216-4402
Thibault Wildi ⓘ https://orcid.org/0000-0002-0193-7132
Jason T. Wright ⓘ https://orcid.org/0000-0001-6160-5888

References

Aigrain, S., Pont, F., & Zucker, S. 2012, MNRAS, 419, 3147
Al Moulla, K., Dumusque, X., Cretignier, M., Zhao, Y., & Valenti, J. A. 2022, A&A, 664, A34
Al Moulla, K., Dumusque, X., Figueira, P., et al. 2023, A&A, 669, A39
Arentoft, T., Kjeldsen, H., Bedding, T. R., et al. 2008, ApJ, 687, 1180
Bedell, M., Hogg, D. W., Foreman-Mackey, D., Montet, B. T., & Luger, R. 2019, AJ, 158, 164
Blackman, R. T., Fischer, D. A., Jurgenson, C. A., et al. 2020, AJ, 159, 238
Boisse, I., Bouchy, F., Hébrard, G., et al. 2011, A&A, 528, A4
Bouchy, F., Bazot, M., Santos, N. C., Vauclair, S., & Sosnowska, D. 2005, A&A, 440, 609
Bouchy, F., Doyon, R., Artigau, É., et al. 2017, Msngr, 169, 21
Cabot, S. H. C., Roettenbacher, R. M., Henry, G. W., et al. 2021, AJ, 161, 26
Cegla, H. M., Watson, C. A., Shelyag, S., et al. 2018, ApJ, 866, 55
Chaplin, W. J., Cegla, H. M., Watson, C. A., Davies, G. R., & Ball, W. H. 2019, AJ, 157, 163
Chaplin, W. J., Houdek, G., Karoff, C., Elsworth, Y., & New, R. 2009, A&A, 500, L21
Claudi, R., Benatti, S., Carleo, I., et al. 2018, Proc. SPIE, 10702, 107020Z
Collier Cameron, A., Ford, E. B., Shahaf, S., et al. 2021, MNRAS, 505, 1699
Collier Cameron, A., Mortier, A., Phillips, D., et al. 2019, MNRAS, 487, 1082
Cosentino, R., Lovis, C., Pepe, F., et al. 2012, Proc. SPIE, 8446, 84461V
Crass, J., Gaudi, B. S., Leifer, S., et al. 2021, arXiv:2107.14291
Cretignier, M., Dumusque, X., Aigrain, S., & Pepe, F. 2023, arXiv:2308.11812
Cretignier, M., Dumusque, X., Allart, R., Pepe, F., & Lovis, C. 2020, A&A, 633, A76
Cretignier, M., Dumusque, X., Hara, N. C., & Pepe, F. 2021, A&A, 653, A43
Davies, G. R., Chaplin, W. J., Elsworth, Y., & Hale, S. J. 2014, MNRAS, 441, 3009
Desort, M., Lagrange, A. M., Galland, F., Udry, S., & Mayor, M. 2007, A&A, 473, 983
Dravins, D. 1982, ARA&A, 20, 61
Dumusque, X. 2018, A&A, 620, A47
Dumusque, X., Cretignier, M., Sosnowska, D., et al. 2021, A&A, 648, A103
Dumusque, X., Pepe, F., Lovis, C., & Latham, D. W. 2015, ApJ, 808, 171
Dumusque, X., Santos, N. C., Udry, S., Lovis, C., & Bonfils, X. 2011a, A&A, 527, A82
Dumusque, X., Udry, S., Lovis, C., Santos, N. C., & Monteiro, M. J. P. F. G. 2011b, A&A, 525, A140
Ervin, T., Halverson, S., Burrows, A., et al. (2022) SolAster: 'Sun-as-a-star' Radial Velocity Variations, Astrophysics Source Code Library, ascl:2207.009
Farret Jentink, C., Mortier, A., Snik, F., et al. 2022, Proc. SPIE, 12182, 1218231
Gilbertson, C., Ford, E. B., Jones, D. E., & Stenning, D. C. 2020, ApJ, 905, 155
Hale, S. J., Howe, R., Chaplin, W. J., Davies, G. R., & Elsworth, Y. P. 2016, SoPh, 291, 1
Halverson, S., Terrien, R., Mahadevan, S., et al. 2016, Proc. SPIE, 9908, 99086P
Harvey, J. W., Hill, F., Hubbard, R. P., et al. 1996, Sci, 272, 1284
Hatzes, A. P. 2002, AN, 323, 392
Huélamo, N., Figueira, P., Bonfils, X., et al. 2008, A&A, 489, L9
Jeffers, S. V., Barnes, J. R., Jones, H., & Pinfield, D. 2013, EPJ Web Conf., 47, 09002
John, A. A., Collier Cameron, A., Faria, J. P., et al. 2023, MNRAS, 525, 1687
Jones, D. E., Stenning, D. C., Ford, E. B., et al. 2017, arXiv:1711.01318
Jurgenson, C., Fischer, D., McCracken, T., et al. 2016, Proc. SPIE, 9908, 99086T
Kjeldsen, H., & Bedding, T. R. 1995, A&A, 293, 87
Kjeldsen, H., Bedding, T. R., Arentoft, T., et al. 2008, ApJ, 682, 1370

Kjeldsen, H., Bedding, T. R., Butler, R. P., et al. 2005, ApJ, 635, 1281
Lafarga, M., Ribas, I., Lovis, C., et al. 2020, A&A, 636, A36
Lanza, A. F., Gizon, L., Zaqarashvili, T. V., Liang, Z. C., & Rodenbeck, K. 2019, A&A, 623, A50
Leet, C., Fischer, D. A., & Valenti, J. A. 2019, AJ, 157, 187
Leite, I., Cabral, A., Abreu, M., & Santos, N. 2022, JPhCS, 2407, 012020
Levine, S. E., Bida, T. A., Chylek, T., et al. 2012, Proc. SPIE, 8444, 844419
Lin, A. S. J., Monson, A., Mahadevan, S., et al. 2022, AJ, 163, 184
Lindegren, L., & Dravins, D. 2003, A&A, 401, 1185
Lovis, C., Dumusque, X., Santos, N. C., et al. 2011, arXiv:1107.5325
Mayor, M., Pepe, F., Queloz, D., et al. 2003, Msngr, 114, 20
Meunier, N. 2021, arXiv:2104.06072
Meunier, N., & Lagrange, A. M. 2019, A&A, 625, L6
Meunier, N., Lagrange, A. M., Borgniet, S., & Rieutord, M. 2015, A&A, 583, A118
Milaković, D., Pasquini, L., Webb, J. K., & Lo Curto, G. 2020, MNRAS, 493, 3997
Molaro, P., Esposito, M., Monai, S., et al. 2013, A&A, 560, A61
Ning, B., Wise, A., Cisewski-Kehe, J., Dodson-Robinson, S., & Fischer, D. 2019, AJ, 158, 210
Park, R. S., Folkner, W. M., Williams, J. G., & Boggs, D. H. 2021, AJ, 161, 105
Pepe, F., Cristiani, S., Rebolo, R., et al. 2013, Msngr, 153, 6
Pepe, F., Mayor, M., Queloz, D., et al. 2004, A&A, 423, 385
Pesnell, W. D., Thompson, B. J., & Chamberlin, P. C. 2012, SoPh, 275, 3
Petersburg, R. R., McCracken, T. M., Eggerman, D., et al. 2018, ApJ, 853, 181

Petersburg, R. R., Ong, J. M. J., Zhao, L. L., et al. 2020, AJ, 159, 187
Phillips, D. F., Glenday, A. G., Dumusque, X., et al. 2016, Proc. SPIE, 9912, 99126Z
Probst, R. A., Lo Curto, G., Avila, G., et al. 2014, Proc. SPIE, 9147, 91471C
Probst, R. A., Milaković, D., Toledo-Padrón, B., et al. 2020, NatAs, 4, 603
Rajpaul, V. M., Aigrain, S., & Buchhave, L. A. 2020, MNRAS, 492, 3960
Rieutord, M., & Rincon, F. 2010, LRSP, 7, 2
Rincon, F., & Rieutord, M. 2018, LRSP, 15, 6
Roettenbacher, R. M., Cabot, S. H. C., Fischer, D. A., et al. 2022, AJ, 163, 19
Saar, S. H. 2003, in ASP Conf. Ser. 294, The Effects of Plage on Precision Radial Velocities, ed. D. Deming & S. Seager (San Francisco, CA: ASP), 65
Saar, S. H., & Donahue, R. A. 1997, ApJ, 485, 319
Santos, N. C., Mortier, A., Faria, J. P., et al. 2014, A&A, 566, A35
Schwab, C., Rakich, A., Gong, Q., et al. 2016, Proc. SPIE, 9908, 99087H
Seifahrt, A., Stürmer, J., Bean, J. L., & Schwab, C. 2018, Proc. SPIE, 10702, 107026D
Strassmeier, K. G., Ilyin, I., Järvinen, A., et al. 2015, AN, 336, 324
Upton, L. A., & Hathaway, D. H. 2023, arXiv:2305.06516
Wilken, T., Curto, G. L., Probst, R. A., et al. 2012, Natur, 485, 611
Wise, A. W., Dodson-Robinson, S. E., Bevenour, K., & Provini, A. 2018, AJ, 156, 180
Zechmeister, M., Anglada-Escudé, G., & Reiners, A. 2014, A&A, 561, A59
Zhao, J., & Tinney, C. G. 2020, MNRAS, 491, 4131
Zhao, L. L., Fischer, D. A., Ford, E. B., et al. 2022, AJ, 163, 171
Zhao, L. L., Hogg, D. W., Bedell, M., & Fischer, D. A. 2021, AJ, 161, 80