



## Three Years of HARPS-N High-Resolution Spectroscopy and Precise Radial Velocity Data for the Sun

Dumusque, X.; Cretignier, M.; Sosnowska, D.; Buchschacher, N.; Lovis, C.; Phillips, D. F.; Pepe, F.; Alesina, F.; Buchhave, L. A.; Burnier, J.

Total number of authors:  
38

Published in:  
Astronomy and Astrophysics

Link to article, DOI:  
[10.1051/0004-6361/202039350](https://doi.org/10.1051/0004-6361/202039350)

Publication date:  
2021

Document Version  
Early version, also known as pre-print

[Link back to DTU Orbit](#)

### Citation (APA):

Dumusque, X., Cretignier, M., Sosnowska, D., Buchschacher, N., Lovis, C., Phillips, D. F., Pepe, F., Alesina, F., Buchhave, L. A., Burnier, J., Cecconi, M., Cegla, H. M., Cloutier, R., Cameron, A. C., Cosentino, R., Ghedina, A., Gonzalez, M., Haywood, R. D., Latham, D. W., ... Wilson, T. G. (2021). Three Years of HARPS-N High-Resolution Spectroscopy and Precise Radial Velocity Data for the Sun. *Astronomy and Astrophysics*, 648, Article A103. <https://doi.org/10.1051/0004-6361/202039350>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Three Years of HARPS-N High-Resolution Spectroscopy and Precise Radial Velocity Data for the Sun

X. Dumusque<sup>1</sup>, M. Cretignier<sup>1</sup>, D. Sosnowska<sup>1</sup>, N. Buchschacher<sup>1</sup>, C. Lovis<sup>1</sup>, D. F. Phillips<sup>2</sup>, F. Pepe<sup>1</sup>, F. Alesina<sup>1</sup>, L. A. Buchhave<sup>14</sup>, J. Burnier<sup>1</sup>, M. Ceconi<sup>7</sup>, H. M. Cegla<sup>1,12</sup>, R. Cloutier<sup>2</sup>, A. Collier Cameron<sup>3</sup>, R. Cosentino<sup>7</sup>, A. Ghedina<sup>7</sup>, M. González<sup>7</sup>, R. D. Haywood<sup>11</sup>, D. W. Latham<sup>2</sup>, M. Lodi<sup>7</sup>, M. López-Morales<sup>2</sup>, J. Maldonado<sup>13</sup>, L. Malavolta<sup>15</sup>, G. Micela<sup>13</sup>, E. Molinari<sup>6</sup>, A. Mortier<sup>4,5</sup>, H. Pérez Ventura<sup>7</sup>, M. Pinamonti<sup>10</sup>, E. Poretti<sup>7,8</sup>, K. Rice<sup>9</sup>, L. Riverol<sup>7</sup>, C. Riverol<sup>7</sup>, J. San Juan<sup>7</sup>, D. Ségransan<sup>1</sup>, A. Sozzetti<sup>10</sup>, S. J. Thompson<sup>4</sup>, S. Udry<sup>1</sup>, and T. G. Wilson<sup>3</sup>

<sup>1</sup> Astronomy Department of the University of Geneva, 51 ch. des Maillettes, 1290 Versoix, Switzerland  
e-mail: xavier.dumusque@unige.ch

<sup>2</sup> Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

<sup>3</sup> SUPA School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews KY16 9SS, UK

<sup>4</sup> Astrophysics Group, Cavendish Laboratory, University of Cambridge, J.J. Thomson Avenue, Cambridge CB3 0HE, UK

<sup>5</sup> Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

<sup>6</sup> INAF - Osservatorio Astronomico di Cagliari, via della Scienza 5, 09047, Selargius, Italy

<sup>7</sup> Fundación Galileo Galilei-IAF, Rambla José Ana Fernández Pérez 7, E-38712 Breña Baja, TF, Spain

<sup>8</sup> INAF - Osservatorio Astronomico di Brera, via E. Bianchi 46, I-23807 Merate (LC), Italy

<sup>9</sup> SUPA, Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, Scotland, UK

<sup>10</sup> INAF—Osservatorio Astrofisico di Torino, Strada Osservatorio 20, Pino Torinese (To) I-10025, Italy

<sup>11</sup> Astrophysics Group, University of Exeter, Exeter EX4 2QL, UK

<sup>12</sup> Department of Physics, University of Warwick, Coventry, CV4 7AL, UK

<sup>13</sup> INAF – Osservatorio Astronomico di Palermo, Piazza del Parlamento 1, 90134 Palermo, Italy

<sup>14</sup> DTU Space, National Space Institute, Technical University of Denmark, Elektrovej 328, DK-2800 Kgs. Lyngby, Denmark

<sup>15</sup> Department of Physics and Astronomy, Università degli Studi di Padova, Vicolo dell'Osservatorio 3, I-35122, Padova, Italy

Received XXX ; accepted XXX

## ABSTRACT

*Context.* The solar telescope connected to HARPS-N has observed the Sun since the summer of 2015. Such high-cadence, long-baseline data set is crucial for understanding spurious radial-velocity signals induced by our Sun and by the instrument.

*Aims.* This manuscript describes the data reduction performed to obtain unprecedented radial-velocity precision for the three years of solar data released along with this paper.

*Methods.* The nearly continuous observation of our Sun has allowed us to detect sub- $m s^{-1}$  systematics in the HARPS-N solar data reduced by the current data reduction software. To improve the radial-velocity precision of the solar data, we reduced them using the new ESPRESSO data reduction software and developed new recipes to mitigate the detected systematics.

*Results.* The most significant improvement brought by the new data reduction is a strong decrease in the day-to-day radial-velocity scatter, from  $1.28$  to  $1.09 m s^{-1}$ ; this is thanks to a more stable method to derive wavelength solutions, but also to the use of calibrations closer in time. We also demonstrate that the current HARPS-N data reduction software induces a long-term drift of  $\sim 1.2 m s^{-1}$ , due to the use of non-stable thorium lines. As a result, the old solar RVs are weakly correlated to the solar magnetic cycle, which is not expected. On the contrary, the newly derived radial velocities are much more correlated, with a Pearson correlation coefficient of  $0.93$ . We also discuss a special correction for ghost contamination, to extract a calcium activity index free from instrumental systematics.

*Conclusions.* Our work leads toward a better understanding of the instrumental and data reduction systematics affecting the HARPS-N spectrograph. The new solar data released, representing an unprecedented time-series of 34550 high-resolution spectra and precise radial velocities will be crucial to understanding stellar activity signals of solar-type stars, with the goal of enabling the detection of other Earths.

**Key words.** Sun: activity – Techniques: radial velocities – Methods: data analysis – Instrumentation: spectrographs – Astronomical data bases – Planets and satellites: detection

## 1. Introduction

The radial-velocity (RV) method is an efficient technique to estimate the mass of exoplanets. If the planets are also transiting, their radius and exact mass can be obtained, thus revealing the exoplanet's density, leading to planetary interior and atmospheric characterisation. The RV technique is also currently the only method for which available instruments reach the sen-

sitivity needed to detect low-mass, non-transiting exoplanets orbiting stars within 40 parsecs from the Sun. Those planets will be prime targets for future direct imaging mission (e.g. [National Academies of Sciences 2018](https://asd.gsfc.nasa.gov/luvoir/resources/)) such as LUVOIR (<https://asd.gsfc.nasa.gov/luvoir/resources/>) and HabEx (<https://www.jpl.nasa.gov/habex/documents/>).

However, when looking for Earth-mass planets orbiting in the habitable zone of solar-type stars, the RV technique is limited

by spurious radial-velocity signals induced by the host star at the  $\text{m s}^{-1}$  level (e.g. Fischer et al. 2016; Plavchan et al. 2018; Blackwood et al. 2020). It is therefore crucial to obtain a better understanding of stellar signals, in order to mitigate them and thus reveal tiny planetary signatures with amplitudes as small as  $10 \text{ cm s}^{-1}$ .

Due to stellar signal timescales ranging from minutes for stellar oscillations to years for magnetic cycles, standard radial-velocity observations obtained either from transit follow-up observations or blind search surveys very often lack the sampling and time baselines necessary to characterise the different types of stellar signals. The lack of good data makes it extremely difficult to understand stellar signals in enough detail, and therefore to find efficient techniques to mitigate them. The ideal solution to make progress is to continuously observe a target which astrophysical parameters and activity behaviour as a function of time are perfectly known, with the same instrument employed for exoplanet characterization. This reflection led to the development of the HARPS-N low-cost solar telescope (LCST, Dumusque et al. 2015). Since then, several solar feeds have been or will be installed at different sub- $\text{m s}^{-1}$  RV precision facilities (HARPS, EXPRES, NEID, MAROON-X, private communication with the different instrument teams). We note that a solar feed is also present on the PEPsi spectrograph (Strassmeier et al. 2015), and on the NIR spectrograph GIANO B (Claudi et al. 2018), however, the data are not dedicated to extreme RV precision.

Since July 2015, HARPS-N has been observing the Sun for several hours nearly every day, with a cadence of 5 minutes, using its solar feed (Phillips et al. 2016). Since then, several papers have analysed the obtained data and investigated solar radial-velocity signals and related activity-index indicators (Collier Cameron et al. 2019; Milbourne et al. 2019; Maldonado et al. 2019; Miklos et al. 2020; Langellier et al. 2020)

The excellent quality of the solar RVs published in Collier Cameron et al. (2019), with instrumental systematics below the  $\text{m s}^{-1}$  level, and the exceptional cadence of these data over several years, revealed instrumental systematics that needed to be understood and corrected for. The first step in this direction was to adapt the publicly available ESPRESSO Data Reduction Software (DRS, Pepe et al. 2020, submitted) to HARPS-N so that the solar data could be reduced with an up-to-date DRS. In this paper, we will refer to the current HARPS-N DRS (version 3.7) as the old DRS and the ESPRESSO pipeline adapted to HARPS-N as the new HARPS-N DRS. Beyond using the new ESPRESSO DRS to extract RV information from raw spectra, we also developed new recipes to account for HARPS-N systematics. We implemented a new procedure to derive wavelength solutions, optimised the selection of thorium lines to improve RV stability on the long-term, used master flat-fields to prevent being photon-noise limited by them when analysing solar data, and corrected ghosts contamination in the calcium activity index time-series.

In Sect. 2, we discuss the HARPS-N solar data set used in this paper and that is part of the data release. In Sect. 3 we discuss the optimisation of the ESPRESSO DRS to HARPS-N and the new recipes implemented to solve for known systematics. We compare and discuss the differences between the RVs obtained with the new and the old DRS in Sect. 4 and we finally conclude in Sect. 5.

## 2. The HARPS-N solar data set and data release

In this paper, and the related data release of the first three years of measurements obtained with the HARPS-N solar telescope, we

only select good quality spectra. As the Sun is resolved with the 3 inch entrance lens of the solar telescope, any kind of absorption, such as clouds, not evenly distributed over the solar disc will induce a huge Rossiter-McLaughlin-type effect, with RVs that can depart from the true value by several hundred  $\text{m s}^{-1}$ , as  $v \sin i$  of the Sun is  $2 \text{ km s}^{-1}$ . To determine if spectra are contaminated by clouds, Collier Cameron et al. (2019) performed a daily regression of  $\log S/N^2$  in echelle order 60 (as a proxy for apparent solar magnitude) against airmass. Outliers from the expected linear extinction relation each day were considered to have been affected by clouds, and were assigned probabilities of belonging to a separate outlier population using a Gaussian mixture model (Foreman-Mackey et al. 2014; Hogg et al. 2010). In the present paper, and the related data release, we only consider the data for which the probability of being of good quality is  $\geq 0.99$ .

With such a threshold, out of the 53417 spectra taken with the solar feed of HARPS-N, 34550 are kept, which corresponds to 65%. The median cadence between consecutive observations, excluding the day-night gap, is 5.42 minutes, which is equal to the effective time between measurements; 5 minute exposure time plus 20 seconds readout. We note that we chose an exposure of 5 minutes to average out the signals induced by p-mode oscillations (e.g. Chaplin et al. 2019).

The data release consists of the following data products:

- The extracted echelle-order spectra, corrected from the instrumental blaze, in the Earth rest-frame. These products are called S2D spectra due to their two dimensional shape. The first and second extensions in the FITS file contain the blaze-corrected extracted flux per pixel and corresponding error for each spectral order. The error corresponds to the photon-noise plus read-out noise of the detector added in quadrature, and divided by the blaze, so that the corresponding error can be directly used with the flux given in the first extension of the FITS file. The third extension corresponds to the quality flag of the pixels for each order, zero being good, and anything else being bad. Hot and bad pixels are flagged that way. Extensions four and five are the wavelength solution in the vacuum and in air, and extensions six and seven are the width of pixels in wavelength in the vacuum and in air, respectively. We note that all wavelengths are in Angströms. Because of dispersion, the size of each pixel in Angströms will change with wavelength, which implies that for a given order, the continuum of an S2D spectrum will show a significant slope. To correct for dispersion and thus get a flat continuum, the easiest way is to divide the flux by the width of pixels in wavelength (divide extension 1 by extension six or seven, depending if you use the wavelength solution in the vacuum or in air).
- The extracted merged spectra, corrected from the instrumental blaze, in the Solar System barycentric rest-frame. These products are called S1D spectra. The only extension in the FITS file includes the wavelength in the vacuum and in air for each point of the merged spectrum, its flux and the quality flag of the point, as defined in the first item above. We note that merged-1d spectra are interpolated on a grid constant in velocity space and not in wavelength space. The step between each point is  $0.82 \text{ m s}^{-1}$ , equivalent to the width of a pixel in velocity space.
- The CCFs obtained by cross-correlating the S2D spectra with a synthetic mask derived from solar observations, thus optimized for the Sun. The first extension in the FITS file gives the CCF measured for each echelle order, with a step

of  $0.82 \text{ m s}^{-1}$ , in addition to the photon-noise weighted average CCF over all orders. Therefore, the extension has the shape  $N_{CCF} \times (N_{ord} + 1)$ , where  $N_{CCF}$  is the number of points of the CCF and  $N_{ord}$  is the number of echelle orders, 69 for HARPS-N. The second extension gives the photon noise errors, and the third one gives the quality flag of each point as defined in the first item above.

In addition to those data products, we also release the time-series of relevant quantities such as the RV, the BIS SPAN, the FWHM, different activity indicators. They are all listed in Table 1 with their descriptions.

The data can be accessed using the Data Analysis Center for Exoplanet (DACE) web-platform at this address <https://dace.unige.ch/sun/>, or using the DACE python API (<https://dace.unige.ch/pythonAPI/> with the solar spectroscopy module). Bulk downloads are available on the main web page, to download all the available products (S2D, S1D and CCF), only the CCF products, or the time-series. Using the solar spectroscopy data base (go to the main web page and then follow the link to the *Solar spectroscopy database*), it is possible to scan through all the released data, select specific observations by using filters (magnifying glass on the right of each column) and download the S2D, S1D or CCF products of specific observations by selecting the rows of interest (click on them), and then using one of the S2D, S1D or CCF buttons on the top right corner of the table. We note that although by default only eight columns are shown on the *Solar spectroscopy database* web page, it is possible to get access to more information by clicking on the gearwheel on the top left corner of the table. All this can also be done using the DACE python API, for people that want to directly access to the data within a python script (<https://dace.unige.ch/pythonAPI/?tutorialId=22>). A small tutorial to perform basic analysis on the solar data released is also available at this address <https://dace-preview.unige.ch/pythonAPI/?tutorialId=23>.

### 3. Deriving precise HARPS-N solar radial-velocity measurements

#### 3.1. The new ESPRESSO DRS

To improve the precision of the HARPS-N solar RVs, we first adapted the ESPRESSO DRS to work with HARPS-N. Although many of the data reduction steps and algorithms implemented in the ESPRESSO DRS find their origins in the HARPS-N DRS, several steps were significantly re-designed and improved with respect to HARPS-N to meet the more challenging requirements of ESPRESSO. A comparison of the old HARPS-N and new ESPRESSO DRS is beyond the scope of this paper, and interested readers are encouraged to read Pepe et al. (2020, submitted) for a general description and in the near future Lovis et al. (2020, in prep.) for all the detail.

#### 3.2. Improved wavelength solution

In the old HARPS-N DRS, a wavelength solution for the science fibre is performed every day by illuminating the corresponding fibre with a thorium-argon lamp. A list of thorium lines, 3692 in total, with their physical position on the detector and their laboratory wavelength is used to detect and measure the centroid of each thorium line in the extracted calibration spectrum. Readers interested in how those lines have been selected should read Lovis & Pepe (2007). Using the line centroids and corresponding

laboratory wavelengths, a wavelength for each pixel is derived by interpolation. For each order individually, this interpolation is performed using a third order polynomial. The old HARPS-N DRS, like the new one, extracts 69 orders from the raw images; thus, daily wavelength solutions are derived using a total of  $4 \times 69 = 276$  free parameters.

The flux throughout an echelle order is not constant due to the blaze response, and is minimal on either side of the order. Therefore, thorium lines in the center of an order are much more constraining for the S/N-weighted polynomial fit used to derive the wavelength solution than thorium lines on both sides of the order. This gives rise to instability in the fitted polynomials and therefore in the wavelength solution from calibration to calibration, which are generally performed every day. This problem was already discussed in several papers (Dumusque 2018; Cersullo et al. 2018; Bauer et al. 2015) and we show in the top row of Fig. 1 an example of how the wavelength solution for a few orders can differ by tens of meter-per-second from one calibration to the next. We show here several wavelength solutions spread over 200 days, and we observe that these strong variations can be seen from one calibration to the next, even when taken only a few minutes apart. This instability in wavelength solutions translates into day-to-day offsets in RV measurements with an estimated root-mean-square (rms) of about  $0.4 \text{ m s}^{-1}$  (Dumusque 2018).

To solve for this instability issue in the wavelength solution, the ESPRESSO DRS uses a combination of a thorium calibration on the science fibre, followed by a Fabry-Pérot calibration on the same fibre to stabilise the wavelength solution. The thorium lines give the absolute wavelength scale, while the Fabry-Pérot's very rich spectrum, with each emission peak being separated by the same distance due to the interferometric nature of the emission, allows us to better constrain the polynomial fit. Readers interested in how to implement such a technique are directed to Cersullo et al. (2018) and Bauer et al. (2015). However, for HARPS-N, such set of calibrations is not present during the entire lifetime of the instrument. As we wanted to adopt a single wavelength solution derivation method for the spectrograph, we decided to implement another strategy, that we call the drift map method. The idea is to use a reference wavelength solution, obtained at a time where consecutive thorium and Fabry-Pérot calibrations on the science fibre are available, and then measure the drift of the thorium lines between each thorium calibration and the reference thorium calibration. A two-dimensional polynomial fit is then performed on the measured drift of the thorium lines and a wavelength solution is obtained for each thorium calibration by shifting accordingly the reference wavelength solution. A detailed description on how to derive a wavelength solution using the drift map method can be found in Appendix A.

To test how the drift map method compares to the thorium-only (old HARPS-N DRS) and combined thorium-Fabry-Pérot techniques when deriving wavelength solutions, we selected 29 thorium calibrations on the science fibre, followed by Fabry-Pérot calibrations on the same fibre (from 2017-03-25 to 2017-11-06). We then derived the wavelength solutions for those calibrations using the thorium-Fabry-Pérot technique and the drift map method, and compared the obtained products to the wavelength solutions derived using the old HARPS-N DRS. In Fig. 1, we show the RV drift, in meters-per-second, between each of these wavelength solutions and the reference wavelength solution used in the drift-map method, obtained on 2017-11-11. From top to bottom, we show the RV drift for the wavelength solution derived using the old HARPS-N DRS, the thorium-Fabry-Pérot combined solution, and the drift map method, respectively. Each column in the figure corresponds to different orders from



Table 1: Description of the time series available for the HARPS-N solar data.

Name	Description	Unit	Type
<i>Filename</i>	Name of the observation	-	STRING
<i>Date BJD</i>	Barycentric Julian Date minus 2400000	days	FLOAT
<i>RV raw</i>	RV of the Sun in the Solar System barycentric rest frame	m s <sup>-1</sup>	FLOAT
<i>RV</i>	RV of the Sun in the heliocentric rest frame, corrected for differential extinction. This value is obtained by subtracting <i>Berv bary to helio</i> and <i>RV diff extinction</i> to <i>RV raw</i>	m s <sup>-1</sup>	FLOAT
<i>RV err</i>	RV error	m s <sup>-1</sup>	FLOAT
<i>Rhk</i>	log( $R'_{HK}$ ) calcium activity index	dex	FLOAT
<i>Rhk err</i>	log( $R'_{HK}$ ) calcium activity index error	dex	FLOAT
<i>Smw</i>	S Mount Wilson calcium activity index	-	FLOAT
<i>Smw err</i>	S Mount Wilson calcium activity index error	-	FLOAT
<i>Bis Span</i>	Bisector span of the CCF	m/s	FLOAT
<i>Bis Span err</i>	Bisector span error of the CCF	m/s	FLOAT
<i>FWHM raw</i>	Raw FWHM of the CCF	m/s	FLOAT
<i>FWHM</i>	FWHM of the CCF, corrected for the solar ecliptic obliquity and Earth orbit eccentricity. See section 3.2 in Collier-Cameron et al. (2019) for more information. We note that in that paper, the authors have to optimize the value of $\gamma$ corresponding to the fraction of the observed $v \sin i$ . They found a value of 1.04 for the old DRS solar data. Performing the same optimisation for the new DRS data lead to a value of 1.15.	m/s	FLOAT
<i>FWHM err</i>	FWHM error of the CCF	m/s	FLOAT
<i>Contrast raw</i>	Raw contrast of the CCF	%	FLOAT
<i>Contrast</i>	Contrast of the CCF, corrected for the FWHM correction so that the equivalent width of the CCF is conserved. See section 3.2 in Collier-Cameron et al. (2019) for more information	%	FLOAT
<i>Contrast err</i>	Contrast error of the CCF	%	FLOAT
<i>Berv</i>	Barycentric Earth RV correction	m/s	FLOAT
<i>Berv bary to helio</i>	Correction to change from the Solar System barycentric to heliocentric rest frame. To change from the heliocentric to barycentric rest frame, just add this term to RV	m/s	FLOAT
<i>RV diff extinction</i>	Estimation of the RV effect induced by differential extinction. See section 2.4 in Collier-Cameron et al. (2019) for more information. To include the effect of differential extinction, just add this value to RV	m/s	FLOAT
<i>Airmass</i>	Sun airmass	-	FLOAT
<i>Coordinates</i>	Sun coordinates	hms / dms	STRING
<i>Texp</i>	Exposure time	sec	FLOAT
<i>Sn order 10</i>	Signal to noise in order 10	-	FLOAT
<i>Sn order 20</i>	Signal to noise in order 20	-	FLOAT
<i>Sn order 30</i>	Signal to noise in order 30	-	FLOAT
<i>Sn order 40</i>	Signal to noise in order 40	-	FLOAT
<i>Sn order 50</i>	Signal to noise in order 50	-	FLOAT
<i>Sn order 60</i>	Signal to noise in order 60	-	FLOAT
<i>Obs quality</i>	Quality flag to assess the observation's quality (no clouds, no calima). Data with values $\geq 0.99$ are all excellent. See section 2.3 in Collier-Cameron et al. (2019) for more information.	-	FLOAT
<i>DRS quality</i>	Quality flag of the data reduction software (true for good, false for bad)	-	BOOLEAN

the blue to the red, including echelle orders 3, 7, 51 and 60. We note that during those 8 months, the instrument drifted by about  $+15 \text{ m s}^{-1}$ , which induces an increasing RV offset when going from the 2017-03-25 to the 2017-11-06 wavelength solutions (purple to yellow colored lines in Fig. 1). However, we measured this instrumental drift on the drift map wavelength solutions and corrected it for all wavelength solutions. To obtain a mean drift value for each drift map wavelength solution, we performed a weighted average taking into account the blaze response of the instrument, and the change of S/N across the echelle orders.

Looking at Fig. 1, it is clear that the thorium only wavelength solutions show more variation from one wavelength solution to the next than the thorium-Fabry-Pérot and drift map wavelength solutions for orders 51 and 60. For blue orders 3 and 7, the thorium-Fabry-Pérot solutions shows similar variations as the thorium only solutions, and only the drift map method gives more stable results. This comes from the fact that the flux on the HARPS-N Fabry-Pérot is rather low in the ten first orders

of the spectrograph. Fabry-Pérot emission peaks on the borders of these blue orders are difficult to detect due to low S/N, and therefore we needed to add the information provided by the thorium lines to improve the wavelength solutions in the blue. Without adding the information from the thorium lines, the thorium-Fabry-Pérot wavelength solutions in the blue part of HARPS-N are even worse than the thorium only solution. We note that the HARPS-N Fabry-Pérot was upgraded after 2017-11-11 and the Fabry-Pérot spectrum now has much more flux in the blue part of the spectrograph, as is also the case for ESPRESSO. For all the orders, from the extreme blue to the extreme red, we see that the wavelength solutions derived with the drift map method are the most stable. We note that although we see that the drift map wavelength solutions present the smallest variations, this is biased by the fact that rather than fitting 276 free parameters to derive a wavelength solution, our drift map model only includes 10 parameters (see Appendix A). In the absence of an absolute wavelength solution derived using a laser frequency comb, it

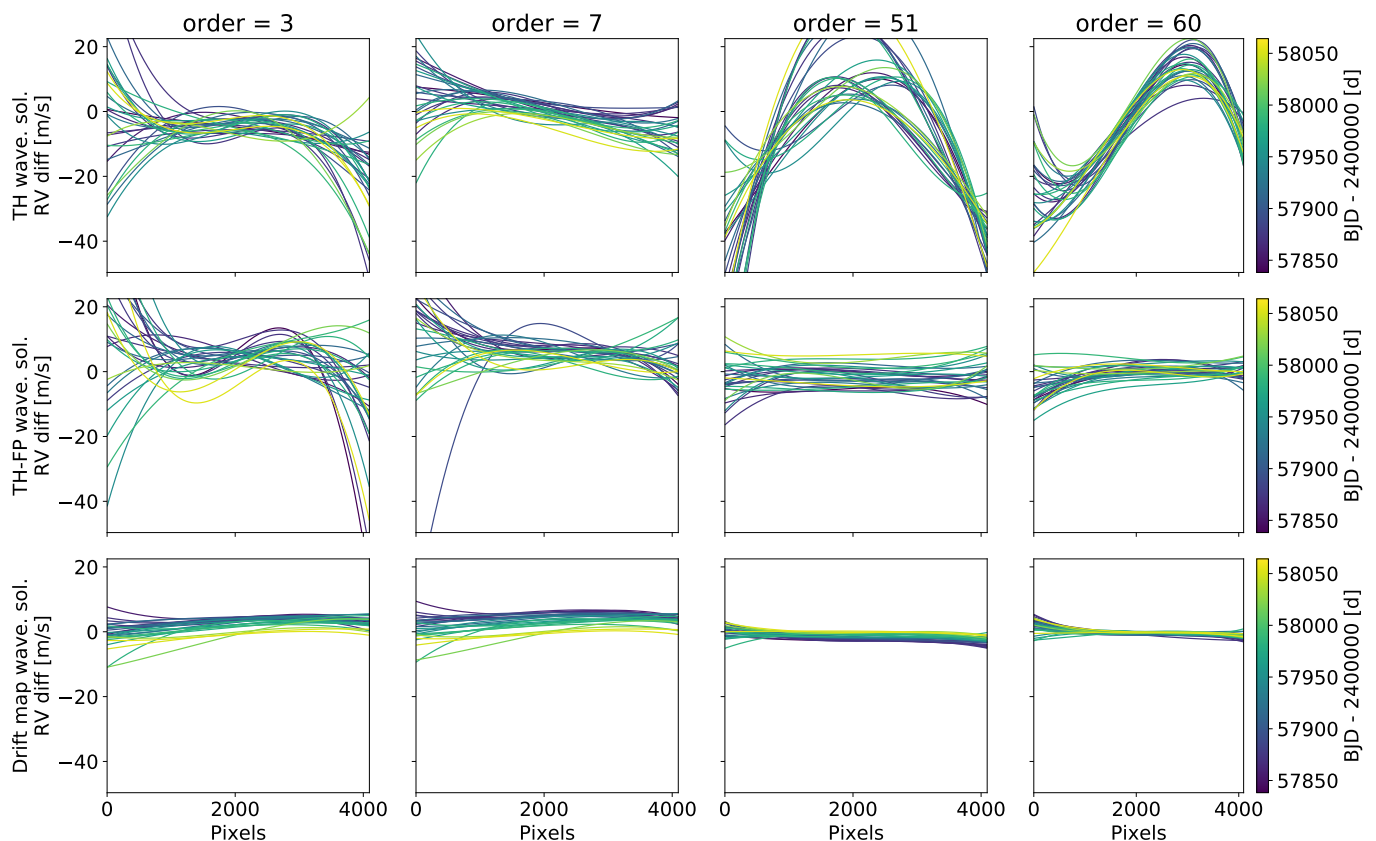


Fig. 1: RV difference between wavelength solutions taken on different days when derived using different algorithms. *Top*: Using only the thorium lines (TH). This is what is done in the old HARPS-N DRS. *Middle*: Using thorium lines and the Fabry-Pérot spectrum (TH-FP). *Bottom*: Using the drift map method presented in this paper, which means using a reference wavelength solution, and shifting it in velocity after measuring the drift of the thorium lines. We note that to obtain the RV difference, we compared all the wavelength solutions with the wavelength solution of reference used in the drift map method, obtained on 2017-11-11.

is impossible to say which of the thorium-Fabry-Pérot or drift map wavelength solutions are the more precise for echelle orders greater than  $\sim 10$ . However, since the drift map method gives for HARPS-N wavelength solutions that are stable even for the bluest echelle orders, we decided to use this technique.

Wavelength solutions set the RV zero-point of the spectrograph each day, and thus directly impact the stability of the measured RVs. To measure the gain in RV stability brought by the new drift map wavelength solutions compared to what was done in the old HARPS-N DRS, we measured the mean RV drift of all wavelength solutions relative to the reference wavelength solution (set for day 2017-11-11, see above). Like in the analysis performed above, the mean RV drift is measured using a weighted average taking into account the blaze response of the instrument and the varying S/N across echelle orders. In the top panel of Fig. 2, we show the result for the old and new wavelength solutions. We clearly see that the old and new DRS gives a similar HARPS-N drift over the three years of the solar data, of about  $140 \text{ m s}^{-1}$ . We also observe sudden jumps at each instrument interventions, either corresponding to a scheduled detector warm-up or to a power failure. In both cases, the observed offset is likely because after warming up and cooling down, the detector does not come back exactly to the previous position (see Sect. 3.6 and Fig. 5 for the rationale behind periodically warming-up the detector).

Over the long-term, several effects can influence the RV stability of the spectrograph, including the ageing of the thorium-argon lamp and the change of the focus of the instrument due to ageing of the optics. In the absence of an absolute calibrator as a laser frequency comb, we can only measure the precision in wavelength solution over short-timescales, considering that the instrument did not drifted. As such, we measured the RV offset between consecutive wavelength solutions that are separated by less than 1.5 days. This selection in time-separation allows to reject the large RV offsets measured between wavelength solutions taken before and after an instrument intervention. In the bottom panel of Fig. 2, we show the distribution of these measured day-to-day RV offsets. As we can see, the new DRS gives a distribution that is more peaked around zero, and that presents less outliers. The MAD of the two distribution is 0.74 and 0.58 for the old and new wavelength solutions, respectively. Consequently, by using the drift map method to derive wavelength solutions, we improve the stability of the RV measurement by a non-negligible  $\sqrt{0.74^2 - 0.58^2} = 0.46 \text{ m s}^{-1}$ . We note that in Dumusque (2018, see appendix C), a similar comparison using HARPS stellar data showed that the drift map wavelength solutions improved the RV precision by  $0.4 \text{ m s}^{-1}$ , which is similar to what is found here.

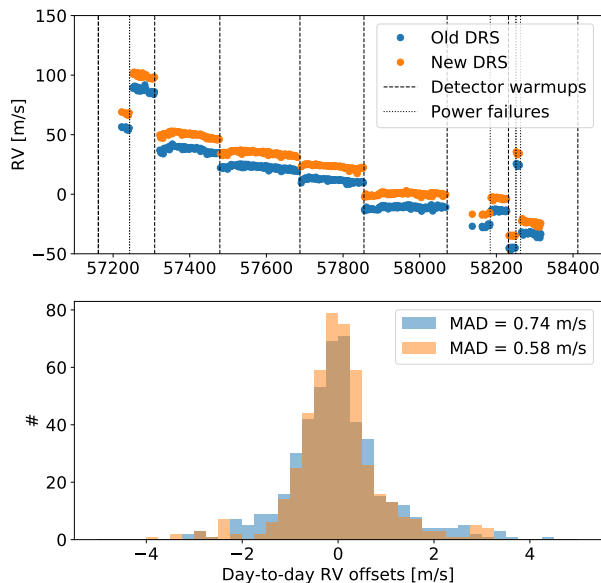


Fig. 2: RV stability of the wavelength solutions derived with the old (blue) and new (orange) DRS. *Top*: RV drift of HARPS-N over-time, measured by computing the RV drift between all wavelength solutions and the wavelength solution of reference (set for day 2017-11-11, see text). We see a general drift over time, interrupted by significant jumps at each warm-up of the detector and power failures (dashed and dotted vertical lines). *Bottom*: Distribution of the RV offsets measured between consecutive wavelength solutions separated by less than 1.5 days, which removes the large offsets observed at each instrument intervention. The MAD of these day-to-day RV offsets is 0.74 and  $0.58 \text{ m s}^{-1}$  for the old and new DRS, respectively.

### 3.3. New set of thorium lines for wavelength solutions

After fitting for the position of the thorium lines, the old and new DRS check if lines are saturated and compare the fitted positions with reference positions. Lines that show a position that is significantly different from the reference are rejected, to prevent biasing the final wavelength solutions. With ageing of the thorium lamp over time, more and more lines are rejected due to lines that start to saturate and lines for which tiny blends become significant and induce a line-profile asymmetry seen as a spurious RV drift. In the left panel of Fig. 3, we can see that out of a total of 3332 thorium lines used by the old HARPS-N DRS on 2014-08-16 (BJD= 2456886), only 2794 (16% less) were still used 5.5 years later<sup>1</sup>, on 2020-01-17 (BJD= 2458866). In addition, we will see in Appendix B that the old HARPS-N DRS was using thorium lines that are not stable over-time. Finally, the old HARPS-N DRS was fitting a combination of several Gaussian profiles on main thorium lines to account for blends, as described in Lovis & Pepe (2007), while the new HARPS-N DRS fits only a single Gaussian profile for each thorium spectral line. All these conditions: a changing set of thorium lines, lines that are not stable over time, and not fitting for blends, can have a significant

<sup>1</sup> We note that for this analysis, we analysed the wavelength solutions obtained for the observation of HD4628, a standard star intensively observed over the years within the HARPS-N Guaranteed Time Observation.

impact on the long-term RV stability, as we will demonstrate below. For all of these reasons, we needed to perform a tailored selection of thorium lines for the new HARPS-N DRS. The description of this selection is rather technical and can be found in Appendix B.

After selecting carefully thorium lines, we are left with 840 lines that can be used to derive wavelength solutions based on the drift map method. Those thorium lines are nearly always used throughout the lifetime of HARPS-N (99.5% of the time, see Appendix B). We note that the entire study performed on the drift map wavelength solutions in Sect. 3.2 uses this new set of thorium lines. In the right panel of Fig. 3, we plotted the RV difference between the wavelength solutions obtained with the old HARPS-N DRS and the new one, which uses the drift map method with the new set of thorium lines. Overall, we clearly see a general trend over time, with a drift of  $\sim 6 \text{ m s}^{-1}$  over 5.5 years, which corresponds to the decreasing number of thorium lines used in the old DRS with time. However, if we look closely, the RV difference is rather stable over the first thousand days of the observations, while it is during this time that the number of thorium lines used decreases the fastest. In fact, the RV difference is very well correlated with the thorium flux ratio (see Fig. B.1), which is the total flux emitted by the thorium lamp at a given time relative to a reference time. The flux ratio is a good proxy to track the ageing of the thorium lamp over time. This demonstrates that the number of thorium lines used is not the only criteria that explains this difference between the old and new wavelength solutions. The RV stability of the lines used also plays an important role (see Appendix B).

To test if this long-term drift observed in the RV difference between the old and new wavelength solutions is introduced by the old HARPS-N DRS, we compared the RVs of several standard stars intensively observed within the HARPS-N Guaranteed Time Observation (GTO) program to search for tiny planetary signals (Motalebi et al. 2015). In the absence of long-term magnetic cycles and long-period planets, we expect the RV of those stars to be flat over time. We can only exclude the presence of strong long-term perturbing signals in this data set, so this analysis has its limitations. With that in mind, we fitted, for each of these stars, a linear drift on the RVs derived using the old and new HARPS-N DRS, between the dates 2014-08-16 and 2020-01-17 (between BJD= 2456886 and BJD= 2458866, same times as the preceding analysis). The measured slopes for the new and old DRS RVs, as well as their differences, are reported in Table 2. For the differences in slope, we obtain rather similar results for all the stars, with values ranging from  $1.34$  to  $2.09 \text{ m s}^{-1} \text{ yr}^{-1}$  and a mean value of  $1.6 \text{ m s}^{-1} \text{ yr}^{-1}$ . This slope is of the same order of magnitude as the  $1.1 \text{ m s}^{-1} \text{ yr}^{-1}$  measured on the RV difference between the old and new wavelength solutions (see right panel of Fig. 3), and therefore points towards the difference in wavelength solutions as being the main origin of the long-term trend difference seen in the stellar RVs. Also, the slopes measured on the new DRS RVs are on average three times smaller than the slopes measured on the old DRS RVs, with values of  $-0.4$  and  $1.2 \text{ m s}^{-1} \text{ yr}^{-1}$ , respectively. Therefore, the new DRS RVs are less affected by long-term instrumental systematics. However, we note that for two stars out of the eleven studied, HD122064 and HD127334, the slope measured on the new DRS is slightly higher. This could be induced by stellar magnetic cycles or low-mass, long-period planets.

The average slope obtained on the new DRS RVs is three times smaller, but is still different from zero, with an average value of  $\sim -0.4 \text{ m s}^{-1} \text{ yr}^{-1}$ . This could be due to astrophysical signals present in the studied RVs. However, out of the 11 stars

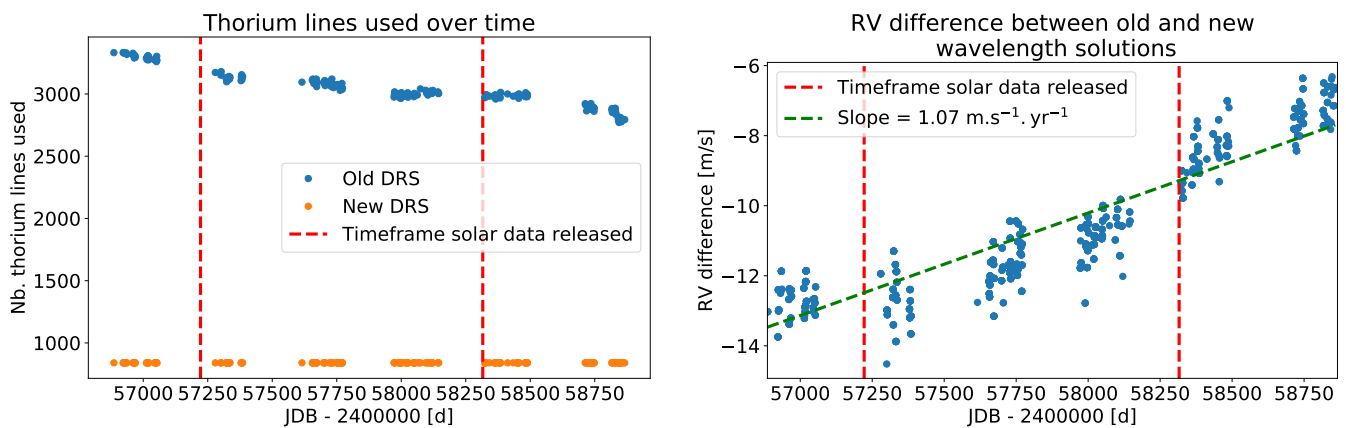


Fig. 3: *Left*: Number of thorium lines used over time to perform the wavelength solution in the old and new HARPS-N DRS. We show here the results for the wavelength solutions used to reduce the observation of HD4628, a standard star intensively followed within the HARPS-N GTO. We see that there is a significant variation in line number used in the old HARPS-N DRS, 16%, while in the new HARPS-N DRS, we always use 840 thorium lines. *Right*: RV difference between the wavelength solutions obtained with the old HARPS-N DRS and the new one, which uses the drift map method with the new set of thorium lines. We can see a long-term trend which is due to the changing number of thorium lines used over time in the old HARPS-N DRS (see Sect. 3.3), as well as considering non-stable thorium lines (see Appendix B.)

studied, 10 show a negative slope in the new RVs, pointing toward a common systematics. This is low number statistics, but it is possible that there is still a remaining long-term systematics in the newly reduced RVs, for which the origin is not explained.

Even with the small discrepancies discussed above, the analysis presented herein is a first demonstration that the new set of thorium lines, combined with the drift map method to derive wavelength solutions, gives more stable RVs on the long-term. We will see in Sect. 4.1 that the new solar RVs better correlate with the calcium activity index, which is expected since RVs are sensitive to stellar magnetic cycles. This gives us a second strong argument in favor of the new RVs being less affected by long-term systematics. As a conclusion to this section, we warn the potential users of the old HARPS-N DRS data that RV measurements are likely affected by a long-term drifts with a slope of  $\sim 1.2 \text{ m s}^{-1} \text{ yr}^{-1}$  between 2014 and 2020. This systematic is induced by the set of thorium lines used to derive the wavelength solutions in the old DRS and should affect all the stars in the same way, independently of their spectral type or magnitude.

### 3.4. Use of afternoon calibrations

In the old HARPS-N DRS, science data are reduced using the closest good quality calibrations back in time. This makes sense for night-time observations, as calibrations are performed in the late afternoon. However, for daily observation of the Sun, this implies that the calibrations used are from the afternoon of the previous day, therefore at minimum more than 12 hours apart. This is not critical for order localisation, flat-fielding (see Sect. 3.5), background and contamination corrections, however, it can be problematic for wavelength solutions, which set the RV zero point of the instrument every day and therefore directly impact RV stability. For each solar observation, with the Fabry-Pérot spectrum on the reference fibre, the wavelength solution is derived by taking the wavelength solution calibration (thorium-argon lamp illuminating the science fibre and Fabry-Pérot spectrum on the reference fibre) and shifting it by the Fabry-Pérot drift observed on the reference fibre. However, this assumes that

Table 2: Comparison between the slope fitted on the RVs derived with the old and the new HARPS-N DRS, for several RV standards observed within the HARPS-N GTO. We note that the linear fit was performed between 2014-08-16 and 2020-01-17 to match the preceding analysis.

Star	Slope old DRS [ $\text{m s}^{-1} \text{ yr}^{-1}$ ]	Slope new DRS [ $\text{m s}^{-1} \text{ yr}^{-1}$ ]	Difference [ $\text{m s}^{-1} \text{ yr}^{-1}$ ]
HD4628	1.34	-0.24	1.58
HD10476	0.79	-0.77	1.56
HD62613	0.90	-0.44	1.34
HD122064	0.67	-0.96	1.64
HD127334	0.53	-0.78	1.31
HD128165	1.82	-0.14	1.96
HD144579	0.89	-0.46	1.36
HD147379	1.83	-0.26	2.09
HD158633	1.25	-0.58	1.83
HD166620	1.22	-0.54	1.76
HD185144	2.01	0.53	1.48
Mean	1.20	-0.42	1.63

the Fabry-Pérot spectrum is stable over time, which is not necessary the case. The first HARPS-N Fabry-Pérot, that was in operation from the beginning of HARPS-N in 2012 to 2017-11-11, was showing internal drifts of a few dozen of  $\text{cm s}^{-1}$  per day on average, with occasional sudden jumps that could be as high as 1 to  $2 \text{ m s}^{-1}$ . This led to its replacement in 2017-11-11 by a more stable system free of instabilities and showing reduced daily drifts. We therefore expect HARPS-N data taken after 2017-11-11 in simultaneous Fabry-Pérot mode to present less instrumental systematics.

The residual internal drift of the Fabry-Pérot implies that, in the used calibration scheme, the wavelength solution is potentially affected by a systematic error that increases with the temporal distance between the calibration exposure and the scientific exposure. While in the old DRS RVs, the data were reduced using the calibration frames of the preceding night, i.e. taken on the preceding afternoon, we reduced all the solar data presented in this paper by using the closest calibrations in the future, which



are the ones obtained after the solar exposures on the late afternoon of the same day. This allows us to better mitigate the systematics induced by the internal drift of the Fabry-Pérot. We did not want to use the closest calibration in time, as morning calibrations, sometimes performed, would induce a jump in the RVs observed within a day, as the wavelength solution calibration used would change from the morning one to the afternoon one around mid-day.

### 3.5. Master flat-fielding

In standard HARPS-N calibrations, five consecutive frames with a tungsten lamp illuminating both science and reference fibres are used to derive the flat field for each echelle order and the corresponding blaze. For most of the night-time observations, even when a few spectra are combined together, those flat-fields have a significantly high S/N to prevent adding noise when spectra are corrected from flat-field during the extraction process. For the Sun, however, as the median number of observations per day is 56, the S/N of the extracted flat field is a limitation when combining spectra reduced with the same calibrations (i.e. spectra taken on the same day). This is, for example, seen in an analysis performed in [Cretignier et al. \(2020\)](#), see Fig.8 and related discussion). We therefore decided to create master flat fields that include more than five raw frames, to increase the S/N and thus be able to work with daily averages of spectra without being significantly affected by the noise present in the flat field products. This implies using frames from different days, which assumes that the instrument flat field does not change significantly over short periods of time.

To test the stability of flat fields over time, we divided all the extracted flat fields by a reference one, in our case HARPN.2012-11-11T00:15:23.151\_FSPECTRUM\_A.fits. We then looked at all the outliers outside of  $1-\sigma$  in the obtained residuals, considering only photon-noise, and measured their average in term of sigma (i.e. value of all outliers in units of  $\sigma$  divided by the number of outliers). This allows us to put more weight on strong outliers. We show this average value of outliers as a function of time in the top panel of Fig. 4. As we can see, this observable varies slowly over time, with sudden jumps that are due to periodic warm-ups of the detector (see Sect. 3.6 for more details about these warm-ups). In between jumps, the variation of the flat field over time is small and we can therefore increase the S/N by accumulating frames over several days. However, we do not want to combine frames on both sides of a jump due to a detector warm-up. Therefore, we used the following strategy to select the frames for a given master flat field. First, we identified jumps as sudden changes of more than 0.2 in average value of outliers. Then, for deriving a master flat field for date 20XX-XX-XX, we looked at all the good raw frames within a time window of  $\pm 20$  days. If there is a jump closer than 20 days from date X, like for example at +10 days, we shift the window accordingly to still have a full window of 40 days. We thus consider the raw frames between dates 20XX-XX-XX-30 days and 20XX-XX-XX+10 days. In this way, we will always keep a high number of frames in master flat fields, with the drawback that master flat fields closer than 20 days from a warm-up will all be identical. This is however not a problem due to the slow variation of the instrument flat field.

The raw frames used for flat fielding on HARPS-N have both fibres illuminated by the tungsten lamp. Therefore, the flat from the science fibre will be contaminated by ghosts from the reference fibre, which are due to secondary reflection of the echelle orders, which will not happen in the case of a science observa-

tion with no flux, the sky, the thorium-argon or the Fabry-Pérot spectrum of the reference fibre. It is thus better to use calibrations where the tungsten lamp is only illuminating one fibre at a time, which is the procedure for ESPRESSO. On HARPS-N, such calibrations exist and are generally used for order localisation, but nothing prevents us from using those raw frames to derive the flat field. The only problem is that there is only two such frames per calibration set (two for the science and two for the reference fibres), compared to the five frames available for standard flat-fielding. However, this can be compensated for by using frames from several different days. In the lower panel of Fig. 4, we show the number of localisation raw frames included in each master flat field. As we can see, the median number of frames is 47, with highest and lowest values at 28 and 73, which is similar the median number (56) of solar frames taken every day. Thus, when using the master flat fields described above to reduce solar data, the noise coming from flat fielding will not significantly affect daily averages of spectra. We note that all the data released with this paper have been reduced using those master flat fields.

### 3.6. Correcting for the calcium activity index contamination

The Mount-Wilson S-index, which is a measure of the chromospheric emission in the core of the Ca II H and K spectral lines, is one of the best activity proxies for solar-type stars. On the Sun, it correlates extremely well with long-term changes in total solar irradiation induced by the presence of magnetically active regions such as faculae and sunspots (e.g. [Lockwood et al. 2007](#), Fig.2). For other stars, this index is used to track the rotational period on the short term (e.g. [Boisse et al. 2009](#)) and magnetic cycles on the long term (e.g. [Wilson 1978](#); [Baliunas et al. 1995](#); [Lovis et al. 2011](#); [Dumusque et al. 2012](#)). To be able to compare the activity of different stars, it is essential to correct the Mount-Wilson S-index from the photospheric flux contaminating the chromospheric emission in the core of the Ca II H and K spectral lines, which is performed when computing the  $\log(R'_{HK})$  activity index ([Noyes et al. 1984](#)). This index is often used to model stellar activity in RV time-series (e.g. [Dumusque et al. 2011](#); [Díaz et al. 2016](#)) or to train non-parametric models like Gaussian Processes to mitigate stellar activity (e.g. [Rajpaul et al. 2015](#)).

However, the S-index or  $\log(R'_{HK})$  is difficult to measure as the related chromospheric emission appears in the extreme blue. At those wavelengths, there is a combination of different physical effects: less stellar spectral energy, more absorption from the Earth's atmosphere, and lower sensitivity of optical fibres, optics and detector, which make it difficult to reach high S/N. In addition, the chromospheric emission happens in the core of the extremely saturated Ca II H and K spectral lines, where the S/N is even lower. Therefore, a careful background correction has to be performed to extract correctly the chromospheric emission. In addition, on HARPS-N, a small leak in the detector cryostat causes humidity to build-up on the detector, which increases its reflectivity over time. A periodic warm-up of the detector solves the issue for a little while. However, with the reflectivity increase, the brightness of ghosts increases as well, modifying the contamination in the extreme blue with time. On the first and second panels of Fig. 5, we show the bluest part of two raw solar science frames of similar S/N obtained with HARPS-N (frames HARPN.2017-04-09T10-06-46.964 and HARPN.2017-04-12T09-09-40.305), with the Sun on the science fibre and the Fabry-Pérot on the reference fibre. The images have been taken just before and after a warm-up of the detector, and show how

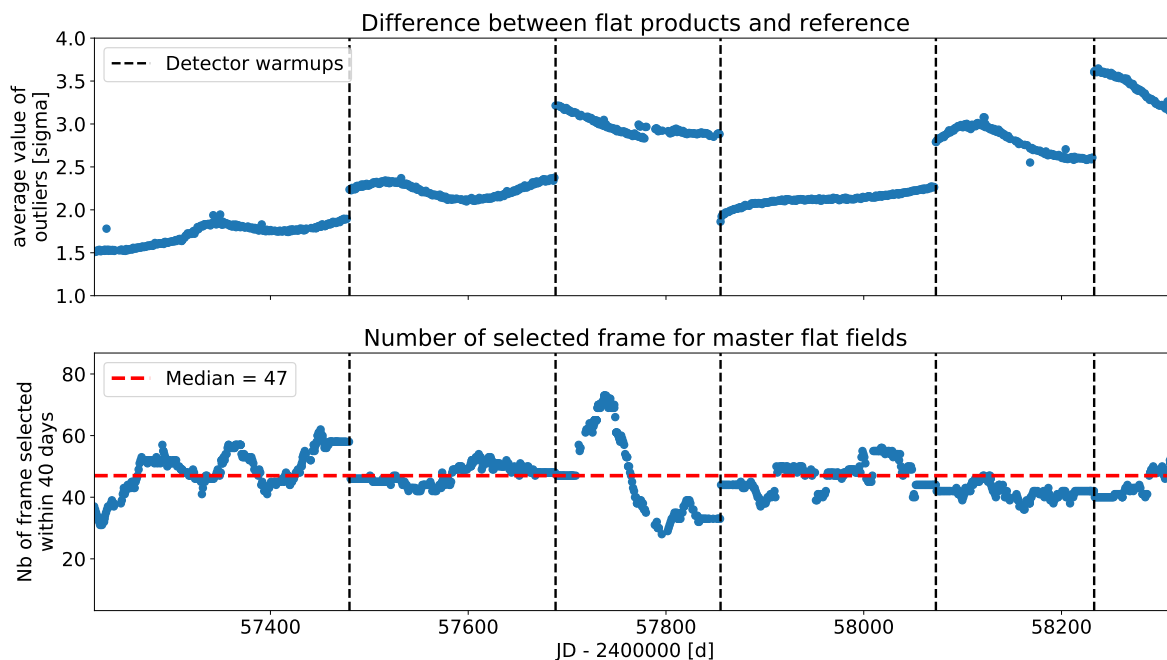


Fig. 4: *Top*: Differences between the flat field over time compared to a reference flat field. We show here the average value of outliers (i.e. the value of all outliers in unit of  $\sigma$  divided by the number of outliers). Sudden jumps can be explained by periodic detector warm-ups. *Bottom*: Number of raw frames used to derive the master flat fields. This number has to be compared to 5, which is the number of raw frames included in the old HARPS-N flat field products, and 56, which is the median number of solar observation per day.

the ghosts significantly reduce in intensity with the intervention.

In Fig. 5, we highlighted the location of the Ca II H and K lines. We note that the core of the Ca II K line, appears on two consecutive orders, extracted orders one and two, and the core of the Ca II H lines appears on the third extracted order. The core of the K line in order one and the H line in order three are significantly affected by ghosts. With humidity building up on the detector over time, ghost contamination increases, and therefore increases the measured  $\log(R'_{HK})$ . After a warm-up, the contamination is significantly reduced, which induces an abrupt offset in the  $\log(R'_{HK})$  time-series. This is clearly visible in the top left panel of Fig. 6, where we show the raw  $\log(R'_{HK})$  extracted by the old HARPS-N DRS along with the time-localisation of each warm-up.

The old and new HARPS-N DRS do not correct for ghost contamination, as this is challenging to perform. First of all, ghosts are at low S/N which makes their extraction difficult. In addition, ghosts are oblique relative to echelle orders and contaminate them over hundreds of pixels. Finally ghosts, which are due to secondary reflection of the echelle orders, do not have a smoothly varying flux, but the flux of the observed solar spectrum. Because it is difficult to know which ghost corresponds to which echelle order, we decided to use a simple technique, described in Appendix C, to correct for the contamination.

The  $\log(R'_{HK})$  measured with and without the correction can be seen in the top left panel of Fig. 6 and the difference between the two is shown in the bottom left panel. It is clear that the increase of flux over time followed by sudden jumps at each detector warm-up is strongly mitigated in the corrected time-series. A closer look at the regions around the third and fourth warm-ups can be seen in the right panel of the same figure. For the

third warm-up, the observed offset without correction is 0.015 dex, and it goes down to 0.005 dex after correction. For the fourth warm-up the offset goes from 0.024 to 0.006 dex. We therefore significantly correct for ghost contamination and the remaining  $\sim 0.005$  dex observed in the residuals at each warm-up is likely due to the simple correction applied in this work. As stated above, a better correction is much more complicated to implement and is out of the scope of the present paper.

The corrected  $\log(R'_{HK})$  activity index that we can see in the top left panel of Fig. 6 has an exquisite precision due to the high S/N solar observations. The median error bar is 0.0014 dex for 5-minute integration times, with a standard deviation of 0.0002 dex. As we can see, the 25-day rotational modulation due to active regions turning with the Sun is clearly visible. Detailed analysis of this  $\log(R'_{HK})$  time-series was performed in Maldonado et al. (2019) and Milbourne et al. (2019). We note that in those papers, the  $\log(R'_{HK})$  time-series were corrected with a similar method as described above.

#### 4. Comparison between the radial velocities reduced with the new and old DRS

In this section, we compare the three years of solar RVs already published in Collier Cameron et al. (2019), obtained with the old HARPS-N DRS and referred to as the old solar RVs, with the RVs obtained with the ESPRESSO DRS and the improvements discussed in this paper (i.e. the new HARPS-N DRS), referred to as the new solar RVs. As described in Sect. 2, we only consider the solar observations for which the quality flag, as defined in Collier Cameron et al. (2019), is  $\geq 0.99$ . This allows us to consider only reliable measurements and reject all observations contaminated by clouds or calima, which can both have a

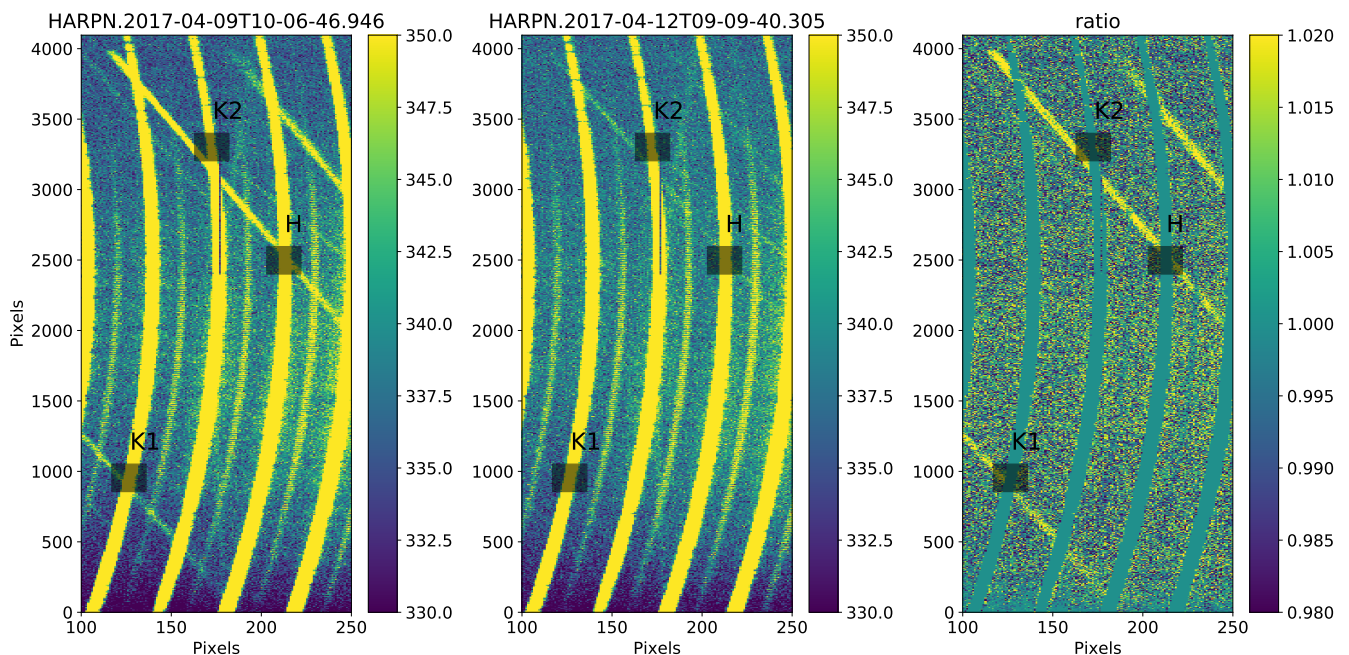


Fig. 5: Raw images from HARPS-N taken before (HARPN.2017-04-09T10-06-46.946, left panel) and after (HARPN.2017-04-12T09-09-40.305, middle panel) a warm-up of the detector, and their ratio (right panel). The location of the core of the Ca II K and H lines in orders one, two and three is highlighted by grey boxes. As we can see, the core of the Ca II K line in order one (K1) and H line in order three (H) are strongly contaminated by ghosts. In addition, we also see that the contamination due to ghosts is significantly reduced with the warm-up process, which removes the humidity accumulated over time on the detector due to a small leak in its cryostat.

significant impact on the RVs. When estimating RV errors, the old HARPS-N pipeline was including photon-noise, drift noise induced by the drift measurement performed on the simultaneous Fabry-Pérot spectrum and the error in wavelength calibration, extracted from the polynomial fit performed to derive it. The new HARPS-N pipeline only includes the first two types of noise, and we clearly demonstrated in Sect. 3.2 that the new wavelength solutions induces day-to-day offsets with a jitter of  $0.58 \text{ m s}^{-1}$ . We note that in the old HARPS-N DRS, the median wavelength solution precision is  $0.35 \text{ m s}^{-1}$ , which is much less than the  $0.74 \text{ m s}^{-1}$  measured in Sect. 3.2. The HARPS-N RVs derived using the old DRS, including the old solar RVs, have thus error bars that are significantly underestimated.

To account properly for the wavelength solution noise, which impact in the same way all the observations performed within the same day, we need to consider a covariance matrix which is block-diagonal, each block regrouping all the observations attached to the same wavelength solution (Delisle et al. 2020). Each block should thus have the same calibration noise of  $0.58 \text{ m s}^{-1}$ . Another way to include correctly this calibration noise is to daily-bin the RV data, and then add in quadrature to the RV errors, the calibration noise of  $0.58 \text{ m s}^{-1}$ .

In the top panel of Fig. 7, we show the comparison between the old and new solar RVs, and the bottom panel of the same figure shows the difference between those two data sets. To see the improvement brought by the new DRS, we divided the time in chunks of 50 days, which corresponds to approximately two solar rotations, and measured the jitter within each chunk using a weighted standard deviation. We find that 17 of the 21 chunks show a greatly reduced jitter in the new solar RVs, with an increased precision of up to 20%. As the jitter is measured over 50 days, this might suggest that the new HARPS-N DRS some-

how mitigates the RV induced by stellar activity on the rotational timescale. However, as we will see in Sect. 4.1, this is not the case. This improvement is mainly brought by the new drift map method with tailored thorium line selection to derive wavelength solutions (see Sect. 3.2), which reduces calibration to calibration RV offsets by  $0.46 \text{ m s}^{-1}$ , and by the use of the afternoon calibrations rather than the ones from the previous day. In Fig. 8, we show the day-to-day RV offsets measured on the old and new RVs after daily-binning the data. As we can see, the new RVs present a distribution with a much smaller MAD of  $1.09 \text{ m s}^{-1}$ , compared to  $1.29 \text{ m s}^{-1}$  for the old RVs. Thus, when analysing day-to-day RV offsets, the new solar RVs show a very significant improvement of  $\sqrt{1.29^2 - 1.09^2} = 0.69 \text{ m s}^{-1}$ .

In the bottom panel of Fig. 7, we can see the difference in RVs between the new and old solar RVs. We see that the two data sets follow each other closely, except for a few points highlighted with red and green vertical lines. After visually inspecting in the old and new DRS RVs if the corresponding data points were following the general RV variation or if they were outliers, we concluded that the 11 red vertical lines corresponds to discrepancies induced by the old DRS. The one green line corresponds to data for which it is difficult to say if the old or the new DRS gives a better solution.

Another significant aspect that can be seen in the RV differences is a long-term drift, which is rather linear. This difference comes from the fact that for the new HARPS-N DRS, we performed a new selection of thorium lines that are used for deriving the wavelength solutions. As explained in Sect. 3.3 and in Appendix B, the new HARPS-N DRS uses a sample of 840 stable thorium lines that are nearly always used throughout HARPS-N history. Using this new selection of lines allows us to stabilise the instrument zero-point over time, and to provide RVs that are less



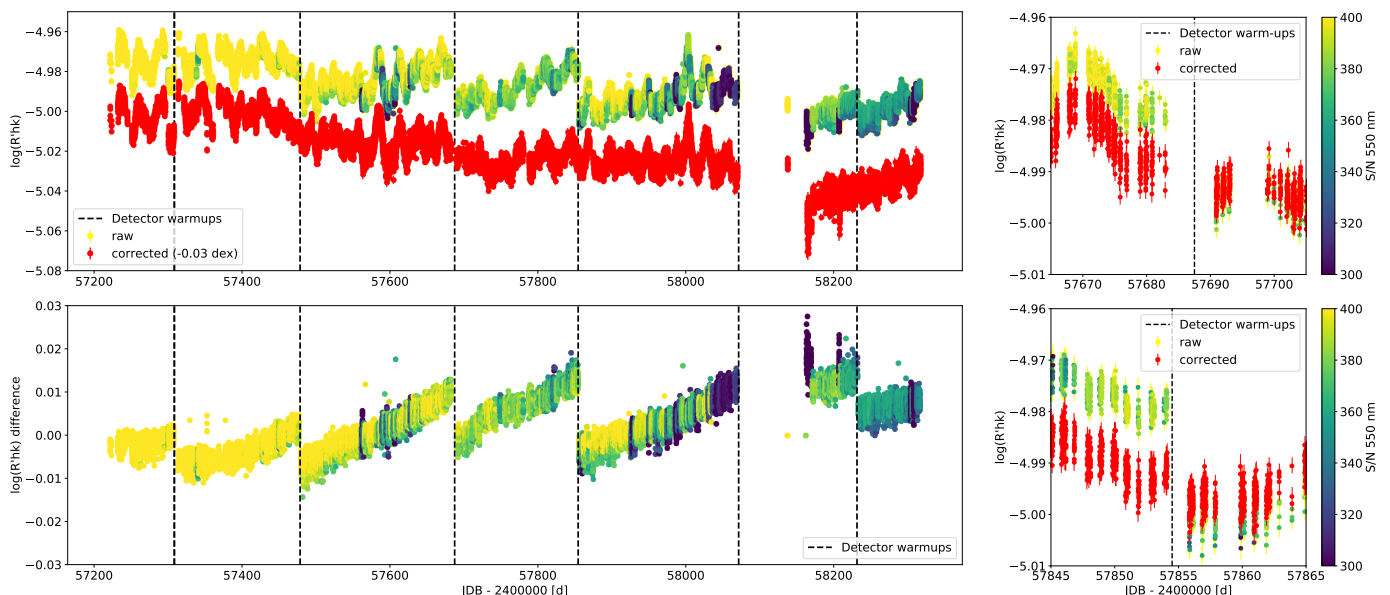


Fig. 6: *Top left*: Comparison between the raw (dots with yellow to blue color scale) and corrected (red dots)  $\log(R'_{HK})$  calcium activity index, as described in this paper. *Bottom left*: Differences between the raw and corrected  $\log(R'_{HK})$  time-series. We clearly see that the slow increase in activity index followed by sudden jumps at each warm-up of the detector seen in the raw  $\log(R'_{HK})$  time-series is strongly reduced after correction. *Top and bottom right*: Closer look at the regions around the third and fourth warm-ups.

affected by long-term systematics (see Table 2). Although the new solar RVs present a long-term drift that is more significant than in the old RVs (see top panel of Fig. 7), after the analysis performed in Sect. 3.3 and the analysis that we will present in Sect. 4.1 we have strong evidences that this drift is indeed real and induced by the solar magnetic cycle.

In Fig. 9, we compared the Generalized Lomb-Scargle periodograms (Zechmeister & Kürster 2009) of the new and old solar RVs. As the new RVs show a significant drift induced by the solar magnetic cycle, we first removed its contribution. To do so, we linearly fitted to the RVs a smooth version of the calcium activity index variation. We note that the smooth version of the calcium activity index was obtained by applying the low-pass filter algorithm described in Press & Rybicki (1989) and by selecting a  $300 \text{ days}^{-1}$  cutoff frequency. We show the raw RVs, as well as the best-fitted linear models for the two data sets in the top panel of Fig. 9. In the bottom panel, we plot the periodograms of the residuals after removing on both data sets the respective long-term trend. At a first glance, the two periodograms are nearly identical. As we can see, the most significant signals can be found around 27 and 13.5 days, the synodic rotational period of the Sun, and its first harmonic, which is expected from stellar activity (e.g. Saar & Donahue 1997; Boisse et al. 2009). In the small inset in the bottom panel of the figure, we clearly see that those signals modulated by stellar rotation have more power in the new DRS RVs, mainly for the peak at 13.4 days. This can be explained by the fact that the new solar RVs are less affected by instrumental and data reduction systematics, and therefore the signal from stellar activity is better characterized. We note however, a slight excess of power around 200 days in the new DRS RVs. This periodicity is also observed in independent observations of the Sun over magnetic cycle 24, estimated using images from the Helioseismic and Magnetic Imager onboard the Solar Dynamics Observatory, in the RV variations as well as in the plage filling factor and unsigned magnetic flux (see

Fig. 5 in Haywood et al. 2020). We conclude that this 200-day periodicity is unlikely to be an instrumental systematic, and that this magnetic activity process is not well traced by the calcium activity index.

#### 4.1. Long-term RV trend induced by the solar magnetic cycle

A strong correlation between the long-term RV and calcium activity index variation for quiet solar-type stars is expected, as the main source of activity for these objects comes from faculae (Meunier et al. 2010b; Haywood et al. 2016; Milbourne et al. 2019). The magnetic field in faculae inhibits locally stellar convection, which suppresses locally the convective blueshift and therefore induces a net redshift (Lindgren & Dravins 2003; Meunier et al. 2010b; Dumusque et al. 2011; Milbourne et al. 2019). From 2015 to 2018, the time-period of the solar data discussed in this paper, magnetic cycle 24 of the Sun was going down towards its minimum level (see Fig. 6). The filling factor of faculae on the solar surface was thus gradually decreasing with time, which increases the overall convective blueshift. It is therefore expected that the values for the solar RVs during this time-period decrease. Looking at Fig. 7, it seems that the new solar RVs are sensitive to the change in solar activity level, which does not seem to be the case for the old RVs.

In Lovis et al. (2011), the authors analysed the RV effect induced by magnetic cycles over a large sample of stars, and derived two empirical relations to predict the RV amplitude induced by the stellar magnetic cycle (see Eq. 9 and 13 from that paper). To apply those relations in our case, we measured the peak-to-peak amplitude of the  $\log(R'_{HK})$  time-series after smoothing it using a 30-day rolling average to remove rotational modulations. We obtained a value of 0.05 dex. Including this value in Eq. 9 and 13 from Lovis et al. (2011), along with an effective temperature of 5780K and metallicity of 0.0, predicts peak-to-peak amplitude in RVs of 8.21 and  $7.74 \text{ m s}^{-1}$ , respec-



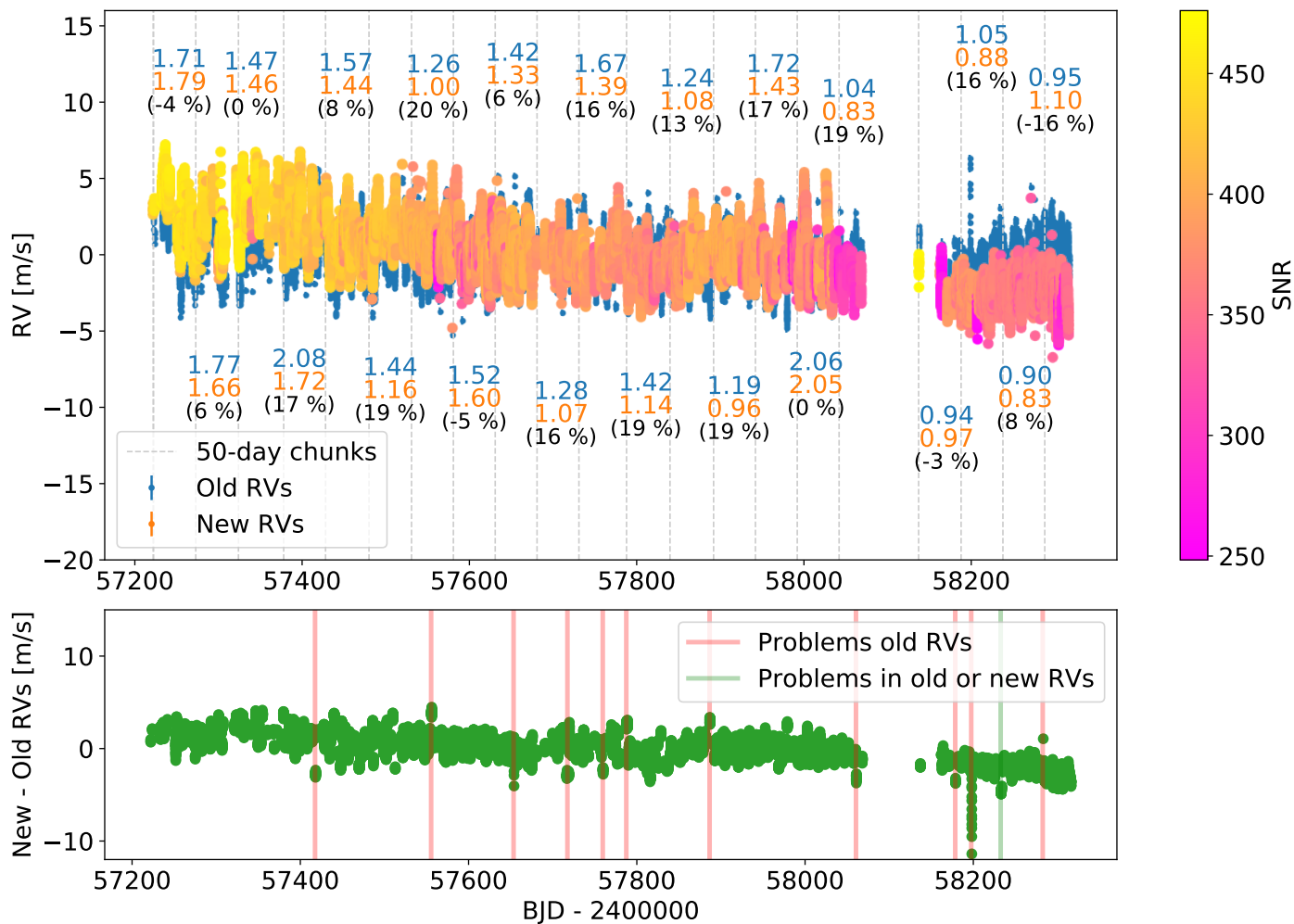


Fig. 7: *Top*: Comparison between the solar RVs reduced using the old HARPS-N and the new HARPS-N DRS. We can see that the data are divided in chunks of 50 days, which corresponds to two solar rotations, and we show for each of them the jitter measured using a weighted standard deviation for the old and new RVs, as well as the improvement in jitter brought by the new RVs in percent. The color scale corresponds to the S/N of the measurements. We see a decrease in S/N over time, likely due to ageing of the Plexiglas dome and/or optical elements along the path of solar light from its entrance into the telescope to its injection into HARPS-N, as the exposure time was always set to 5 minutes. *Bottom*: Difference between the new and old solar RVs, for which a long-term trend is clearly visible. We highlighted with red and green vertical lines the significant discrepancies in those RV differences. The red lines corresponds to problems in the old DRS RVs, while for the green vertical line, it is not clear if the correct RVs are the ones derived with the old or new DRS. We note that around BJD=2458100 (December 2017), the fiber guiding the light from the solar telescope to HARPS-N was broken following an intervention in the HARPS-N calibration unit. This led to observations with very low S/N, giving RVs that are not precise enough. Those data were rejected from the data released, explaining the absence of data around BJD=2458100.

tively. These values have to be compared with the peak-to-peak amplitude measured on the long-term variations seen in the old and new RVs (blue and orange lines in the top panel of Fig. 9), which are  $3.8$  and  $7.6 \text{ m s}^{-1}$ , respectively. The peak-to-peak RV amplitude measured on the new solar RVs matches the prediction of Lovis et al. (2011), therefore it is likely those RVs better represent the reality.

Looking at the correlation between the RV and the  $\log(R'_{HK})$  time-series gives an additional argument in favour of the new solar RVs being a better representation of the reality. In Fig. 10, we show on the top plot the correlation between the RVs and the  $\log(R'_{HK})$  for long-term variations, which in this case means

that the time-series has been smoothed using a 30-day rolling average. The bottom plot shows the short-term variation, meaning the raw time-series minus the smoothed ones. As we can see, the long-term RV variations correlates much better with the  $\log(R'_{HK})$  for the new solar RVs, with a strong correlation of  $R = 0.93$  compared to  $R = 0.72$  for the old RVs. Regarding the short-term variations, the new solar RVs do slightly better, with  $R = 0.53$  compared to  $R = 0.48$ , however, the difference is not very significant. This analysis shows clearly what is already known regarding the correlation between RV and the calcium activity index: on the timescale of a magnetic cycle, the correlation is extremely strong, while on the timescale of rotation, a correla-

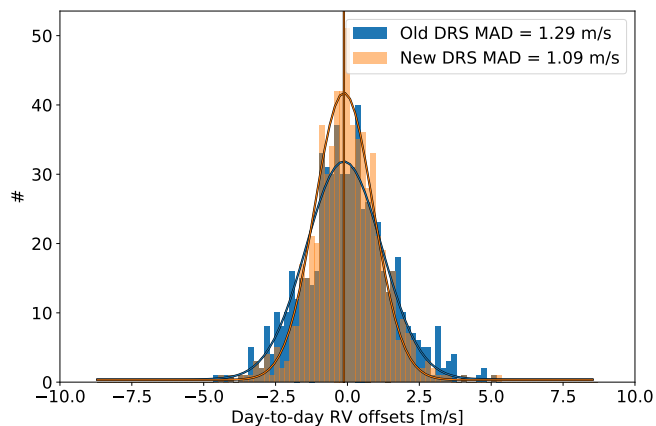


Fig. 8: Day-to-day RV offsets measured on the old and new solar RVs after a daily-binning of the data.

tion is present, however weaker. Therefore, the calcium activity index is not the best proxy to account for the RV stellar activity signals on the rotational timescale. The unsigned magnetic flux seems to be a better proxy (Haywood et al. 2020).

With the two arguments given above, we infer that the stronger long-term drift observed in the new solar RVs is real, and is induced by the solar activity cycle. When fitting the new DRS RVs with a linear drift, we obtain a slope of  $-1.7 \text{ m s}^{-1} \text{ yr}^{-1}$ . Considering the analysis performed in Sect. 3.3 where we estimated a slope difference between the new and old DRS RVs of  $1.6 \text{ m s}^{-1} \text{ yr}^{-1}$ , we understand why the old DRS solar RVs are rather flat over time. The long-term RV drift induced by the magnetic cycle is nearly perfectly compensated for by the long-term RV systematics induced by the old DRS wavelength solutions.

## 5. Conclusion

In this paper, we described the data reduction that allowed us to obtain unprecedented RV precision for the three years of HARPS-N solar data that are released alongside this manuscript (see Sect. 2 for more information about the released products). We also compared the newly derived RVs with the ones already published in Collier Cameron et al. (2019), and clearly demonstrated that the new data are less affected by instrumental and data reduction systematics.

Compared to the Collier Cameron et al. (2019) RV data that were reduced using the old DRS, we reduced the new ones using the ESPRESSO DRS tailored to HARPS-N (see Sect. 3.1). In addition, we developed a new method to derive wavelength solutions (see Sect. 3.2) and curated the list of thorium lines that are used when deriving a wavelength solution, so that only the most stable lines over time are used (see Sect. 3.3 and Appendix B). We also reduced the data taking the closest calibrations in the future rather than in the past (see Sect. 3.4) to mitigate any systematics induced by Fabry-Pérot internal drifts, and we used master flat fields that have a much higher S/N than standard flat fields, to prevent the solar observations being limited by them (see Sect. 3.5). We also extracted a calcium activity index, which has been corrected from ghost contamination (see Sect. 3.6).

We performed several analysis in Sect. 4 to compare the new and old DRS RVs. When dividing the solar data in chunks of 50 days to still see the 25-day effect of stellar rotation, we observe that for most of the chunks, the new solar RVs are less affected

by noise, with improvement of up to 20% (see Fig. 7). As the main source of jitter in the data is induced by stellar activity on the rotational timescale, we also compared the new and old RV data set on the timescale of a day, which is more affected by instrumental and data reduction systematics. On that timescale, it is clear from Fig. 8 that the new RVs are much less affected by instrumental and data reduction systematics as the measured jitter goes down to 1.09, compared to  $1.29 \text{ m s}^{-1}$ . As systematics adds up in quadrature, this implies that the new reduction reduces a noise of  $0.69 \text{ m s}^{-1}$  present in the old solar RVs. This improvement is mainly brought by the new wavelengths solution derivation method, the optimal selection of thorium lines, and the use of calibrations closer in time.

We also observed that the new RVs show a stronger drift with time than the old RVs. However, since this long-term drift correlates much better with the stellar magnetic cycle,  $R = 0.93$  for the new RVs and only 0.72 for the old ones (see top panel Fig. 10), we concluded that the observed drift in the new RVs should be real (see Sect. 4.1). This drift being the manifestation of the solar magnetic cycle in RV measurements, which is expected (e.g. Dumusque et al. 2011; Meunier et al. 2010a; Lindgren & Dravins 2003). We note that the old solar RVs are rather flat with time, because the effect induced by the magnetic cycle is compensated for by a long-term data reduction systematics induced by the use of thorium lines that are not stable over-time and that are not always used. This is something that is demonstrated on different stellar observations of standard stars at the end of Sect. 3.3 and related Table 2. Users analysing old DRS RVs should be cautious about the interpretation of long-term drift, as most of the standard stars studied in this paper showed a linear drift with time, with a slope of  $\sim 1.2 \text{ m s}^{-1} \text{ yr}^{-1}$ .

From the different analysis performed in this paper, it is clear that the new solar RVs and corresponding high-resolution spectra and cross-correlation functions (CCFs) should be preferred compared to the old DRS products, as the instrumental and data reduction jitter is smaller, but also because those data better reflect the effect from the solar magnetic cycle. After releasing these data through the DACE web-platform (see end of Sect. 2), the next effort from the HARPS-N collaboration will be to reduce all the HARPS-N data using this new version of the DRS, as the gain in term of precision and stability is clearly demonstrated on the Sun.

On a final note, we reduced at best all the observed systematics on the HARPS-N solar data set. The release of this unprecedented time-series of spectra and radial velocities is crucial to improve our understanding of stellar signals affecting the Sun and solar-type stars. We hope that the community will use such data to develop novel methods to mitigate stellar signals in radial-velocity data sets, with the goal of enabling the detection of other Earths.

*Acknowledgements.* XD is grateful to The Branco Weiss Fellowship–Society in Science for its financial support. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 851555). ACC acknowledges support from the Science and Technology Facilities Council (STFC) consolidated grant number ST/R000824/1. This work has been carried out in the framework of the National Centre for Competence in Research PlanetS supported by the Swiss National Science Foundation (SNSF). AM acknowledges support from the senior Kavli Institute Fellowships. This research used the DACE platform developed in the frame of PlanetS (<https://dace.unige.ch>). This work was performed under contract with the California Institute of Technology (Caltech)/Jet Propulsion Laboratory (JPL) funded by NASA through the Sagan Fellowship Program executed by the NASA Exoplanet Science Institute (RDH).

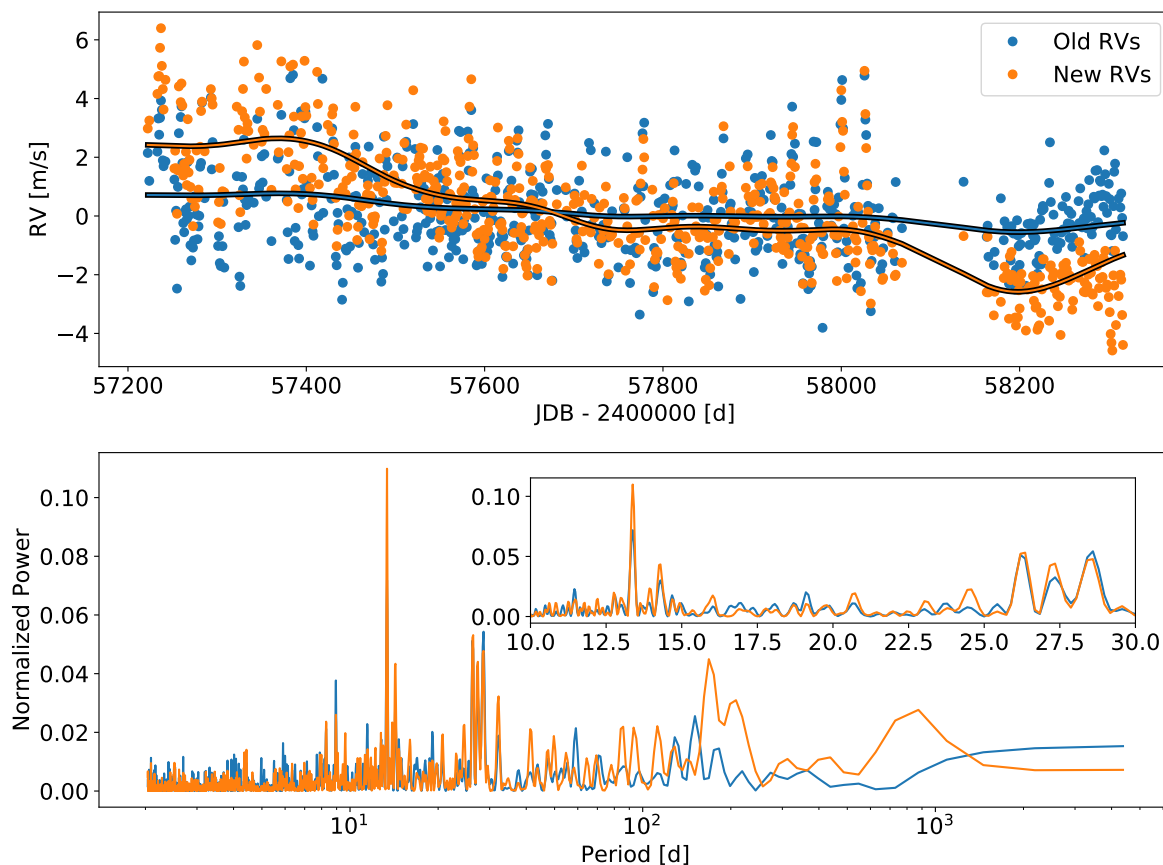


Fig. 9: *Top*: Raw RVs derived with the old and new HARPS-N DRS. The blue and orange lines correspond to a smooth version of the calcium activity index variation fitted to the RVs using a scaling factor, to remove the contribution from the solar magnetic cycle. *Bottom*: Generalized Lomb-Scargle Periodogram of the residuals after removing the long-term trend observed in the RVs. The small inset shows a zoom around the periods 10 to 30 days.

## References

- Baliunas, S. L., Donahue, R. A., Soon, W. H., et al. 1995, *ApJ*, 438, 269
- Bauer, F. F., Zechmeister, M., & Reiners, A. 2015, *A&A*, 581, A117
- Blackwood, G., Gaudi, B. S., Burt, J., et al. 2020, in *American Astronomical Society Meeting Abstracts*, American Astronomical Society Meeting Abstracts, 374.01
- Boisse, I., Moutou, C., Vidal-Madjar, A., et al. 2009, *A&A*, 495, 959
- Cersullo, M. F., Coffinet, A., Chazelas, B., Lovis, C., & Pepe, F. 2018, *A&A*, submitted
- Chaplin, W. J., Cegla, H. M., Watson, C. A., Davies, G. R., & Ball, W. H. 2019, *AJ*, 157, 163
- Claudi, R., Ghedina, A., Pace, E., et al. 2018, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 10700, *Ground-based and Airborne Telescopes VII*, 107004N
- Collier Cameron, A., Mortier, A., Phillips, D., et al. 2019, *MNRAS*, 487, 1082
- Cretignier, M., Dumusque, X., Allart, R., Pepe, F., & Lovis, C. 2020, *A&A*, 633, A76
- Delisle, J. B., Hara, N., & Ségransan, D. 2020, *A&A*, 638, A95
- Díaz, R. F., Ségransan, D., Udry, S., et al. 2016, *A&A*, 585, A134
- Dumusque, X. 2018, *A&A*, 620, A47
- Dumusque, X., Glenday, A., Phillips, D. F., et al. 2015, *ApJ*, 814, L21
- Dumusque, X., Lovis, C., Ségransan, D., et al. 2011, *A&A*, 535, A55
- Dumusque, X., Pepe, F., Lovis, C., et al. 2012, *Nature*, 491, 207
- Fischer, D., Anglada-Escude, G., Arriagada, P., et al. 2016, *ArXiv e-prints* [arXiv:1602.07939]
- Foreman-Mackey, D., Hogg, D. W., & Morton, T. D. 2014, *ApJ*, 795, 64
- Haywood, R. D., Collier Cameron, A., Unruh, Y. C., et al. 2016, *MNRAS*, 457, 3637
- Haywood, R. D., Milbourne, T. W., Saar, S. H., et al. 2020, *arXiv e-prints*, arXiv:2005.13386
- Hogg, D. W., Bovy, J., & Lang, D. 2010, *arXiv e-prints*, arXiv:1008.4686
- Langellier, N., Milbourne, T. W., Phillips, D. F., et al. 2020, *arXiv e-prints*, arXiv:2008.05970
- Lindgren, L. & Dravins, D. 2003, *A&A*, 401, 1185
- Lockwood, G. W., Skiff, B. A., Henry, G. W., et al. 2007, *ApJS*, 171, 260
- Lovis, C., Dumusque, X., Santos, N. C., et al. 2011, *ArXiv e-prints* [arXiv:1107.5325]
- Lovis, C. & Pepe, F. 2007, *A&A*, 468, 1115
- Maldonado, J., Phillips, D. F., Dumusque, X., et al. 2019, *A&A*, 627, A118
- Meunier, N., Desort, M., & Lagrange, A.-M. 2010a, *A&A*, 512, A39
- Meunier, N., Lagrange, A., & Desort, M. 2010b, *A&A*, 519, A66+
- Miklos, M., Milbourne, T. W., Haywood, R. D., et al. 2020, *ApJ*, 888, 117
- Milbourne, T. W., Haywood, R. D., Phillips, D. F., et al. 2019, *ApJ*, 874, 107
- Motalebi, F., Udry, S., Gillon, M., et al. 2015, *A&A*, 584, A72
- National Academies of Sciences, Engineering, M. 2018, in *Exoplanet science strategy (Consensus study report)*, Washington, DC: The National Academies Press.
- Noyes, R. W., Hartmann, L. W., Baliunas, S. L., Duncan, D. K., & Vaughan, A. H. 1984, *ApJ*, 279, 763
- Phillips, D. F., Glenday, A. G., Dumusque, X., et al. 2016, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 9912, *Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation II*, 99126Z
- Plavchan, P., Cale, B., Newman, P., et al. 2018, *arXiv e-prints*, arXiv:1803.03960
- Press, W. H. & Rybicki, G. B. 1989, *ApJ*, 338, 277
- Rajpaul, V., Aigrain, S., Osborne, M. A., Reece, S., & Roberts, S. 2015, *MNRAS*, 452, 2269
- Saar, S. H. & Donahue, R. A. 1997, *ApJ*, 485, 319
- Strassmeier, K. G., Ilyin, I., Järvinen, A., et al. 2015, *ArXiv e-prints* [arXiv:1505.06492]
- Vaughan, A. H., Preston, G. W., & Wilson, O. C. 1978, *PASP*, 90, 267
- Wilson, O. C. 1978, *ApJ*, 226, 379
- Zechmeister, M. & Kürster, M. 2009, *A&A*, 496, 577

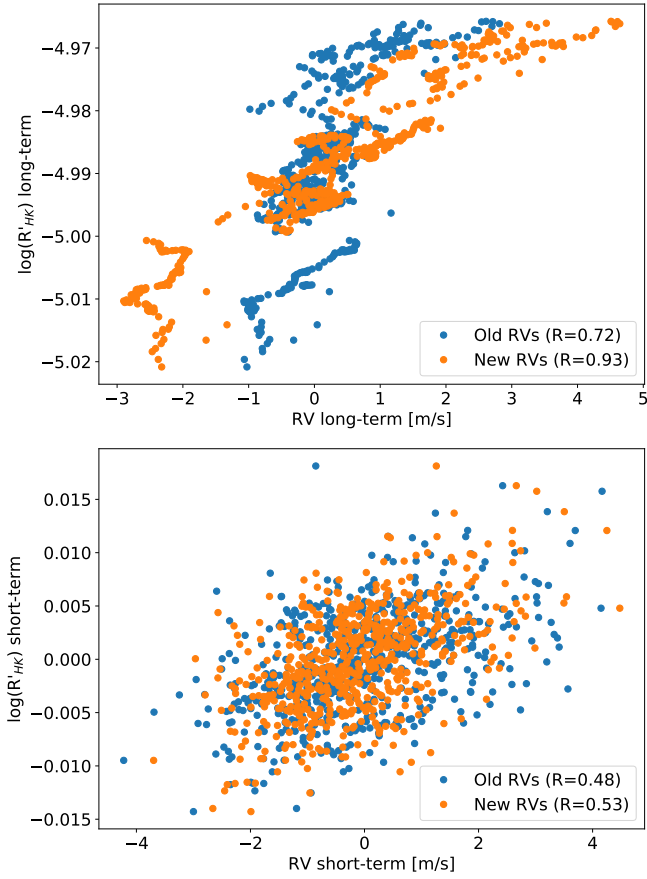


Fig. 10: *Top*: Correlation between the RVs and the  $\log(R'_{HK})$  for long-term variations, which in this case means that the time-series have been smoothed using a 30-day rolling average. *Bottom*: Correlation between the short-term variations, meaning the raw time-series minus the smoothed ones. We show in the legend the values for the R Pearson correlation coefficient for the new and old solar RVs.



## Appendix A: Wavelength solution derived using the drift map method

In Fig. A.1 we show the different steps performed to derive and validate the quality of a wavelength solution using the drift map method. The first step is to measure the drift of the thorium lines between a thorium calibration, for which we want a wavelength solution, and a reference thorium calibration, for which a reference wavelength solution already exists. For this solar data release, we derived a reference wavelength solution for the day 2017-11-11, by combining information from a thorium calibration followed by several consecutive Fabry-Pérot calibrations, which allowed us to obtain a very high-S/N Fabry-Pérot spectrum, and therefore a precise reference wavelength solution. We then fit on the measured drift of all thorium lines, a smooth two-dimensional polynomial of shape  $\Delta X = \sum_{i,j} X^i O^j$ , where  $i + j \leq 3$ ,  $X$  corresponds to pixel position in the dispersion direction, and  $O$  to order number, to account for variations in the cross-dispersion direction. The corresponding wavelength solution is obtained by shifting the reference wavelength solution using the fitted 10 free parameters two dimensional polynomial. We note that we selected a two dimensional polynomial with  $i + j \leq 3$ , as tests using polynomials of smaller or higher dimensions ( $i + j \leq 2$  or  $i + j \leq 4$ ) were giving residuals with a larger scatter or no significant improvement, respectively.

## Appendix B: Tailored selection of thorium lines for the new HARPS-N DRS

Because in the new way of deriving wavelength solutions, described in Sect. 3.2 and Appendix A, we only use a small number of free parameters, we can significantly reduce the number of thorium lines to use, as long as we are left with thorium lines evenly-distributed on the detector. We therefore decided to remove thorium lines significantly affected by blends. This decision is driven by the fact that with the ageing of thorium lamps, it is difficult to predict the behavior of blended thorium lines. If the flux ratio between a blend and the main thorium line it affects changes with the ageing of the lamp, the resulting variation in asymmetry will be associated to a change in line position, which is not necessarily the case. This spurious change in line position will affect the derived wavelength solution, and therefore the RV stability of the spectrograph on the long-term.

To detect thorium lines significantly affected by blends, we compared for one thorium calibration frame the position of the 3332 thorium lines fitted with the old and new DRS. The old DRS fits multiple Gaussian components to a thorium line if blends are known (Lovis & Pepe 2007). The new DRS only fits for a single Gaussian component. Therefore, if a thorium line is significantly blended, we expect to see a large difference in the derived pixel position between the old and new DRS. Therefore, we compared the pixel position derived with the old and new DRS for the 3692 thorium lines used by the old DRS, and rejected all the thorium lines for which the difference in position, expressed in velocity, was larger than  $200 \text{ m s}^{-1}$ . We also rejected lines for which the expected amplitude was smaller than 3000, to consider only strong emission lines. After this selection, we were left with a total of 2781 thorium lines.

To reject even further blended thorium lines, and lines that would not be properly fitted by the new DRS, we visually inspected the 2781 thorium lines left, and kept only the ones that had a Gaussian shape without any significant blend, and no sign of other emission lines within seven pixels from the line centre, as the new HARPS-N DRS fits a single Gaussian profile for each

thorium line in a window of fourteen pixels. After this manual selection, we were left with 1366 thorium lines.

To reject lines that were not stable over time, we then reduced all the HARPS-N thorium calibrations with the obtained line list and kept only the ones that were used 99.5% of the time<sup>2</sup>. At the end of this process, we were left with 1138 thorium lines.

We then compared, for the thorium calibration used for the reference wavelength solution (set for day 2017-11-11, see Sect. 3.2), the fitted position of the thorium lines between the science and the reference fibres, and removed the lines for which the difference from the median was more than four times the MAD. This process rejected an additional 108 lines, therefore, providing a number of 1030 good thorium lines.

Looking at the drift over time of each individual thorium line left after our drastic selection, we still saw that some lines were showing a significant linear drift over time of dozens of  $\text{m s}^{-1}$  per year, and some others were strongly correlated with the thorium flux ratio, which is the total flux emitted by the thorium lamp at a given time relative to a reference time. The flux ratio is a good proxy to track the ageing of the thorium lamp over time. In the left panel of Fig. B.1, we show the thorium flux ratio of the unique thorium lamp used on HARPS-N. As we can see, the thorium flux ratio changes significantly over time, mainly after BJD=2458300, where we observe an exponential drift, characteristic of a thorium lamp in the end of its life. In the right panel of Fig. B.1, we show the RV drift of a thorium line that shows a strong correlation of  $R=0.92$  with thorium flux ratio. This is likely explained by this thorium line being affected by a blend, for which the flux ratio relative to the main thorium line changes with the ageing of the lamp. It is clear that such a thorium line shows a strong systematics over time and therefore, should not be used.

To assess the stability of the remaining 1030 thorium lines in our selection, we measured three different quantities. For each lines, we fitted a model corresponding to a linear drift plus a linear correlation with thorium flux ratio. We then considered the slope of the linear drift and the Pearson correlation coefficient between the RV drift residuals after removing the fitted slope and the thorium flux ratio as two relevant quantities of line stability. We also measured the ratio between the rms of the RV drift residuals and the median error in RV, which corresponds to our third quantity. This last quantity probes the presence of remaining signals in the RV drift residuals. We measured those quantities for all the thorium lines used in the old and new DRS, and reported their values in Fig. B.2. As we can see in this figure, the new selection of thorium lines described above (in orange) removes a large fraction of lines showing significant slopes and large Pearson correlation coefficients. Regarding the ratio between the RV drift residual rms over the RV error, we see that the values obtained for the old DRS is significantly smaller than for the new DRS. This is because the old DRS is giving larger errors in RV when fitting for the thorium line centroids. We clearly see in the bottom right panel of Fig. B.1 that the error bars for the old DRS are twice as large as the ones for the new DRS. This can be explained by the fact that the old DRS is fitting multiple Gaussian components on the profile of a thorium line, to account for blends, which increases the estimated RV errors.

Looking at the different distributions obtained in Fig. B.2 for the new DRS thorium line selection, we decided to reject

<sup>2</sup> We note that we wanted to remove all the lines that were not used 100% of the time, however, a very few bad calibrations were rejecting nearly all the thorium lines. To reject those bad calibrations from our analysis, we therefore lowered slightly the acceptance threshold.



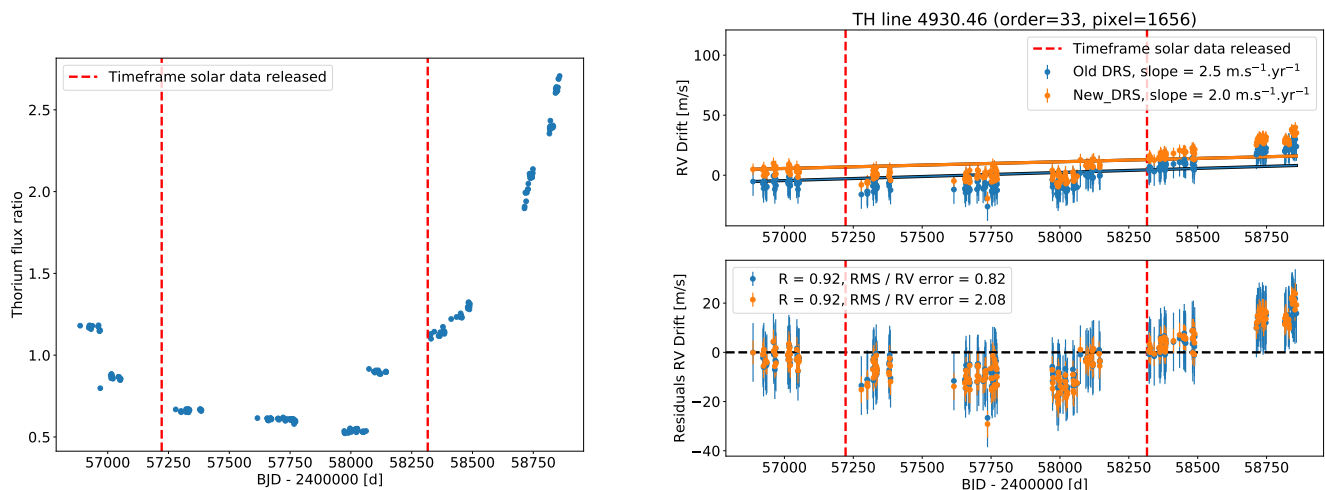


Fig. B.1: *Left*: Thorium flux ratio over time, which is the total flux emitted by the thorium lamp at a given time relative to a reference time. *Top right*: RV drift of the thorium line at 4930.46 Å measured with the old (blue) and new (orange) DRS. On both data set, we fitted a model corresponding to a linear drift plus a linear correlation with the thorium flux ratio. We report in the legend the measured slope of the linear drift. *Bottom right*: RV drift residuals after subtracting the fitted linear drift. We report in the legend the R Pearson correlation coefficient between those residual and the thorium flux ratio, as well as the ratio between the rms of those residual and the median error in RV.

Table B.1: Thorium line list used by the new HARPS-N DRS. Wavelengths are in the vacuum and are coming from Lovis & Pepe (2007). This table only shows a small fraction of the data that are available in electronic format.

Wavelength [Å]	Element	HARPS-N order	HARPS-N pixel
3880.743712	TH1	0	462
3886.869574	TH2	0	985
3891.008506	TH1	0	1351
3892.828887	TH1	0	1514
3894.755139	TH1	0	1690
3901.983250	TH2	0	2371
3902.767225	TH1	0	2447
...	...	...	...
6814.655533	TH2	67	2890
6839.384102	TH1	68	235
6841.181722	TH1	68	319
6856.401613	TH2	68	1051
6857.581950	TH1	68	1109
6884.710078	TH1	68	2526
6885.214016	TH1	68	2554

constrain the two-dimension polynomial fit performed to derive wavelength solutions (see Appendix A). As we can see in Fig. B.3 the final selection of thorium lines is evenly distributed over the HARPS-N detector. The final selection of thorium line that is used with the new DRS is composed of 840 lines, that are reported in Table B.1.

## Appendix C: Correcting the calcium activity index from ghost contamination

To calculate the contamination induced by ghosts on the measured flux in the core of the Ca II H and K lines, we optimally extracted the flux, using the profile of the corresponding orders, on both sides of the echelle orders containing the Ca II H and K lines. To select on which pixel to extract the flux, we used the order localisation, and shifted the localisation by plus or mi-

nus eight pixels in the cross dispersion direction. We highlighted the localisation of those pixels as blue and orange lines in the left panel of Fig. C.1. As the separation between the science and calibration fibre is at least sixteen pixels, these eight pixel shift allows to extract the flux from ghosts, without being significantly contaminated by the solar or Fabry-Pérot spectra. Optimally extracting the flux using the profile of the orders allows us to measure the exact flux that contaminates the science fibre. We describe in detail in the caption of Fig. C.1 how the ghost contamination is measured in the core of the Ca II H line. We perform the same measurements for the core of the Ca II K line appearing in order one and two of HARPS-N, and finally correct from this contamination before measuring the corrected S-index, and relative  $\log(R'_{HK})$ .

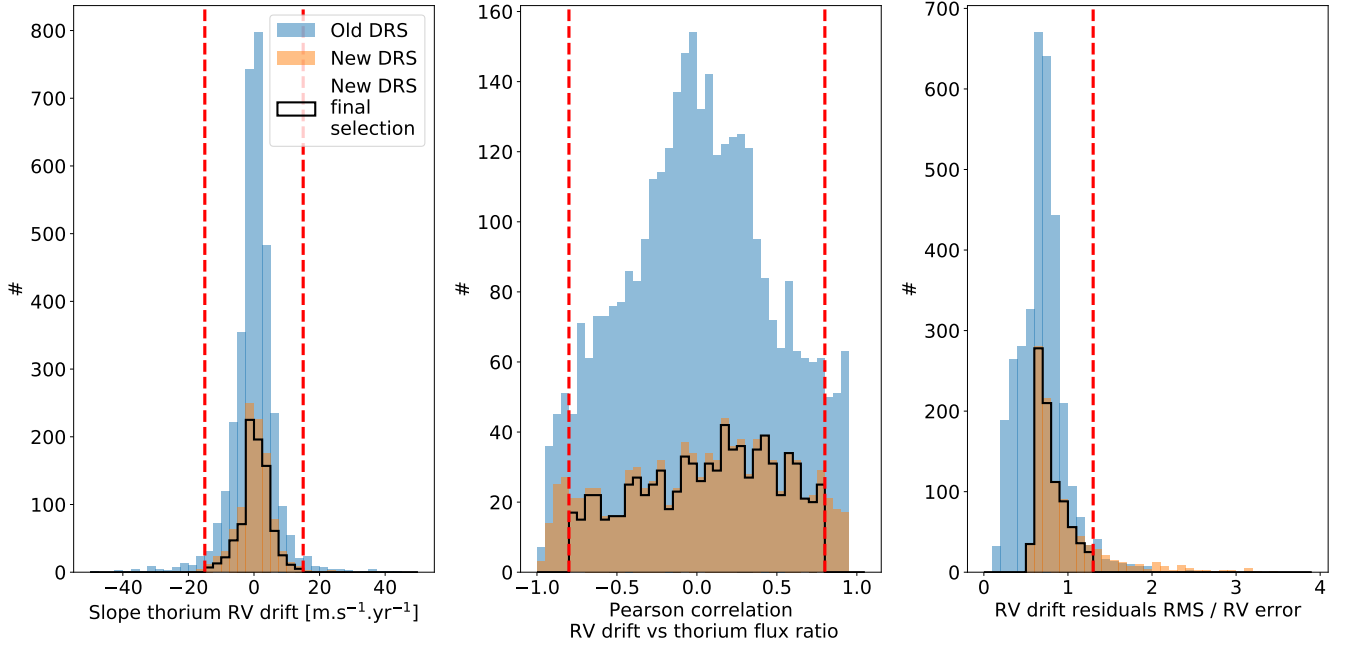


Fig. B.2: *Left*: Slope measured on the RV drift of all the thorium lines used in the old (blue) and new (orange) DRS. *Middle*: Pearson correlation coefficient between the RV drift of each thorium line and the thorium flux ratio (see Fig. B.1). *Right*: RMS of the RV drift residuals of all thorium lines when subtracting a linear drift, divided by the median error in RV. For each subplot, the red vertical dashed lines correspond to the cutoffs used to reject thorium lines showing strong systematics ( $\pm 15 \text{ m.s}^{-1}.\text{yr}^{-1}$  for slope,  $< 0.8$  for Pearson correlation coefficient and  $< 1.2$  for residual rms over error in RV). The black histogram corresponds to the final selection of lines used in the new DRS, after rejecting lines using the different cutoffs.

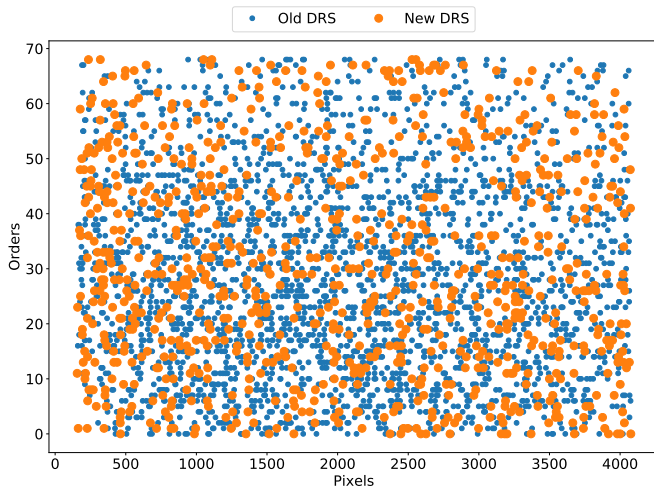


Fig. B.3: Position of the thorium lines used by the old (blue) and new (orange) DRS as a function of their position on the detector. We note that the position in dispersion direction is in pixel, and the position in cross-dispersion direction in echelle order number.



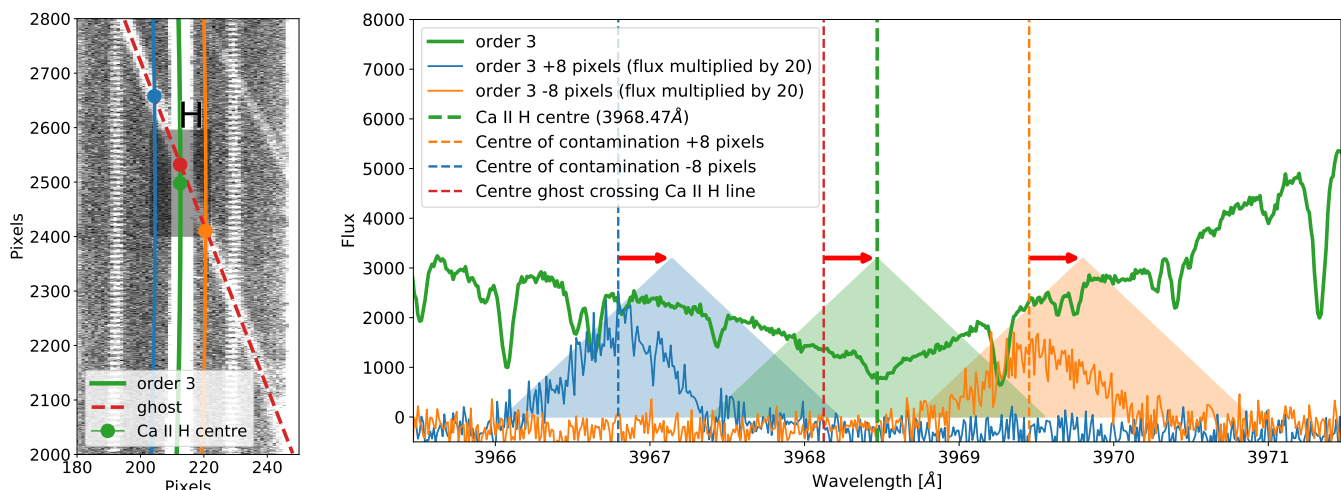


Fig. C.1: Derivation of the chromospheric emission in the core of the Ca II H line with estimation of the contamination induced by the perturbing ghost. *Left*: Raw image centred on extracted order three with the position of the Ca II H line core. *Right*: The extracted spectra at different positions. The green spectrum is the stellar spectrum extracted from order three (green line on the left panel). The blue and orange spectra are the flux optimally extracted, using the same order profile as order three, minus and plus eight pixels in the cross-dispersion direction on each side of order three (blue and orange lines on the left panel). We note that the extracted fluxes for those spectra have been multiplied by 20 to highlight the contamination from the ghost (red dashed line on the left panel). The vertical green dashed line corresponds to the centre of the Ca II H line core (green dot on the left panel), and the blue and orange vertical dashed lines to the centre of the ghost on both sides of order three (blue and orange dots on the left panel). Assuming that the ghost does not show any curvature, which is a good approximation from one side of the order to the other, we measure the crossing of the ghost and order three (red dot on left panel, and red vertical dashed line on the right panel) as the middle between the orange and blue dots. To measure the chromospheric emission in the core of the Ca II H, the extracted flux from the green spectrum is multiplied with the triangular green response of width  $1.09 \text{ \AA}$  (Vaughan et al. 1978). To measure the effective ghost contamination, we multiply the contaminating flux on the blue and the orange spectra with the same triangular response, taking into account that the ghost does not cross order three at the center of the Ca II H line core (the difference is highlighted by the red arrows). Finally, the contamination inside the core of the Ca II H line is taken as the average of the contaminating flux on both sides of the orders.