

# Precise and efficient modeling of stellar-activity-affected solar spectra using SOAP-GPU

Yinan Zhao<sup>1</sup>, Xavier Dumusque<sup>1</sup>, Michael Cretignier<sup>2</sup>, Khaled Al Moulla<sup>1</sup>, Momo Ellwarth<sup>3</sup>, Ansgar Reiners<sup>3</sup>, and Alessandro Sozzetti<sup>4</sup>

<sup>1</sup> Department of Astronomy of the University of Geneva, 51 chemin de Pegasi, 1290 Versoix, Switzerland

<sup>2</sup> Department of Physics, University of Oxford, OX13RH Oxford, UK

<sup>3</sup> Georg-August Universität Göttingen, Institut für Astrophysik und Geophysik, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany

<sup>4</sup> INAF - Osservatorio Astrofisico di Torino, Via Osservatorio 20, I-10025 Pino Torinese, Italy  
e-mail: zhaoyinan2121@gmail.com

December 19, 2024

## ABSTRACT

*Context.* One of the main obstacles in exoplanet detection when using the radial velocity (RV) technique is the presence of stellar activity signal induced by magnetic regions. As the most advanced techniques to mitigate this signal are reaching a level better than one meter per second, it is difficult to evaluate their performance: instrumental systematics start to be similar in magnitude, and therefore it is impossible to know the ground truth of the stellar activity signal. In this context, a realistic simulated dataset that can provide photometry and spectroscopic outputs is needed for method development.

*Aims.* The goal of this paper is to describe two realistic simulations of solar activity obtained from SOAP-GPU and to compare them with real data obtained from the HARPS-N solar telescope. For this purpose, both simulated spectral time series cover the time window of HARPS-N solar observation, but nothing prevents SOAP-GPU from modeling the data over different time spans.

*Methods.* We describe two different methods of modeling solar activity using SOAP-GPU. The first models the evolution of active regions based on the spot number as a function of time. Other physical parameters are either drawn from observed solar distributions or modeled with empirical relations. The second method relies on the extraction of active regions from the Solar Dynamics Observatory (SDO) data. The location of spots and faculae on the solar disk at each timestamp are derived from the magnetogram and intensity maps and are fed into SOAP-GPU to simulate the corresponding spectra.

*Results.* The simulated spectral time series generated with the first method shows a long-term RV behavior similar to that seen in the HARPS-N solar observations. The effect of stellar activity induced by stellar rotation is also well modeled with prominent periodicities at the stellar rotation period and its first harmonic. The comparison between the simulated spectral time series generated using SDO images and the HARPS-N solar spectra shows that SOAP-GPU can precisely model the RV time series of the Sun to a precision better than 0.9 m/s. By studying the width and depth variations of each spectral line in the HARPS-N solar and SOAP-GPU data, we find a strong correlation between the observation and the simulation for strong spectral lines, therefore supporting the modeling of the stellar activity effect at the spectral level. The correlations are weaker for shallow lines, although it is likely that their lower signal-to-noise ratio does not allow a meaningful comparison.

*Conclusions.* We introduce two methods for modeling solar activity using SOAP-GPU. With only sunspot numbers as input, we accurately capture the long-term magnetic cycle and rotational features. Additionally, we effectively model shift and depth variations at the spectral line level by using data from SDO. These simulated solar spectral time series serve as a useful test bed for evaluating spectral-level stellar activity mitigation techniques.

**Key words.** Methods: data analysis – Techniques: radial velocities – Techniques: spectroscopic – Stars: activity

## 1. Introduction

The radial velocity technique was the first method used to detect a planet orbiting a solar-type star, 51 Peg b (Mayor & Queloz 1995), and it remains the most efficient method today for measuring planetary minimum masses. However, even considering the impressive progress that was made over the past 30 years, mainly in terms of instrumentation, this method is still limited by the stellar signal, and mainly its activity component, which prevents the detection of true Earth analogs (e.g., Crass et al. 2021). In recent years, the field has moved toward probing and mitigating stellar activity in the spectral (or cross-correlation function) domain, where the stellar signal, inducing an asymmetry of spectral lines and affecting each of them differently, should be easier to disentangle from a planetary signal that simply shifts all spec-

tral lines in wavelength (e.g., Feng et al. 2017; Dumusque 2018; Cretignier et al. 2020; Collier Cameron et al. 2021; Cretignier et al. 2021; Al Moulla et al. 2022; de Beurs et al. 2022; Cretignier et al. 2022; Zhao et al. 2022; Al Moulla et al. 2024).

In the quest for the detection of Earth-analogs, developing stellar activity correction techniques that enable the mitigation of this perturbing signal to a few dozen cm/s requires exquisite datasets. In that direction, it is now common for ultra-stable spectrographs used for exoplanet detection during the night-time to also observe the Sun during the day. The following instruments all have solar input and provide high-resolution spectra ( $R \sim 100\,000$ ) for which a radial-velocity (RV) precision better than 1 m/s is reached: HARPS-N (Dumusque et al. 2015; Phillips et al. 2016; Collier Cameron et al. 2019; Dumusque et al. 2021), HARPS (Al Moulla et al. 2023), EXPRES (Zhao et al. 2022),

NEID (Lin et al. 2022), KPF (Rubenzahl et al. 2023). These solar data are an ideal benchmark for developing and testing the performance of different mitigating techniques for stellar activity (e.g., Collier Cameron et al. 2021; de Beurs et al. 2022; Langellier et al. 2021; Zhao et al. 2024). Although these solar observations are the closest to night-time measurements, a comparison of several solar datasets shows that at the sub-meter-per-second level, some differences are visible, likely due to instrument systematics that are not perfectly calibrated (e.g., Zhao et al. 2023). In addition, even though the solar datasets now include more spectra than had been obtained on other single stars, the limited number of spectra available<sup>1</sup> makes it challenging to start exploring sophisticated machine learning techniques such as deep learning. For these reasons (contamination with instrumental signals and a limited amount of data), having a realistic simulation of stellar activity is crucial in order to develop and test the efficiency of mitigation techniques.

Several methods have been proposed to model solar activity to test the detection sensitivity of Earth-like planets or to derive solar activity proxies. Meunier et al. (2010) were likely the first to realistically model the RV of the Sun including the inhibition of convective blueshift (CB) in active regions, and demonstrated that for solar-type slow rotators this effect dominates over the flux effect induced by the contrast difference of active regions. Borgniet et al. (2015) modeled solar cycle 23 in radial velocity (RV) by parameterizing the active region properties as observed on the Sun. Haywood et al. (2016) used the better observations provided by the Solar Dynamics Observatory (SDO), compared to the work of Meunier et al. (2010) that used SOHO, to model the solar RVs. They used the dopplergrams, intensity continuum, and magnetogram images from the Helioseismic and Magnetic Imager (HMI/SDO, Schou et al. 2012) to derive the RV components including the CB inhibition and the flux imbalance seen in facula and spot regions. All these simulations are providing realistic RVs affected by stellar activity for the Sun, however, without modeling properly the solar spectra. The first such an attempt was likely the work performed by Gilbertson et al. (2020) that used a modified version of SOAP 2.0 (Dumusque et al. 2014), but the simulation is likely too simplistic as it only parameterizes the spot region configurations without considering faculae, which are the main contributors of stellar signal for the Sun. In addition, SOAP 2.0 is rather limited by computational performance in the number of spectra that can be modeled.

Recently, the SOAP-GPU code was developed to lift the SOAP 2.0 constraint on computational effort thanks to GPU performance (Zhao & Dumusque 2023). In addition to being more efficient than SOAP 2.0, SOAP-GPU also includes more physics based on recent solar observations obtained by Löhner-Böttcher et al. (2019) and Ellwarth et al. (2023). The simulation now includes realistic center-to-limb variation (CLV) of the quiet Sun spectrum. Such observations at different limb-angles, which cover the full visible spectrum, still do not exist for active regions. Therefore SOAP-GPU takes as input the CLV observed in Cavallini et al. (1985) for the FeI line at 6301.5008Å, and propagates the observed bisector to all spectral lines. Thanks to SOAP-GPU, it is now possible to model the effects of solar activity at the spectral level with likely the best precision possible.

<sup>1</sup> The HARPS-N solar telescope, the first of its kind on sky, started operation in the summer of 2015. We now have accumulated ~ 200000 spectra; however, many measurements are obtained each day. Within a day, the activity of the star which evolves on the timescale of the solar rotation, does not change significantly. Therefore it is the total number of days with observations that counts if we want to probe stellar activity. In the case of HARPS-N, this number equals approximately 1500 days.

In this paper we present two simulated solar spectral time series datasets modeled with SOAP-GPU. The first, described in Sect. 2, uses as input the solar spot number time series to model solar activity based on the known properties of solar active regions. The second, described in Sect. 3, uses as input the SDO images to extract at each time the location and size of active regions. As this is as close as possible to reality, we then compare the modeled spectra with the ones derived from the HARPS-N solar telescope. Finally, we discuss and draw our conclusions in Sect. 4. The simulated datasets and the simulation code are publicly available on Zenodo and Github.<sup>2</sup>

## 2. Solar spectral time series modeling using spot number

Two main factors are taken into account when using SOAP-GPU to simulate activity-affected spectral time series, i) the properties of active regions and ii) the properties of input spectra for both the quiet and active regions. The properties of active regions, both in the spatial and the temporal domain, are well described in Borgniet et al. (2015) and we present in Sect. 2.1 the parametrization we used as input of SOAP-GPU to model solar activity. In Sect. 2.2 we describe the input spectra that are used in SOAP-GPU to model the quiet and active regions. Finally, the simulated spectral time series is presented in Sect. 2.3.

Throughout this paper, the differential rotation of the solar disk is included according to the equation  $\omega = \omega_0 + \omega_1 \sin^2(\theta)$ , where  $\omega_0 = 14.371^\circ/\text{day}$  and  $\omega_1 = -2.587^\circ/\text{day}$  (Borgniet et al. 2015; Zhao & Dumusque 2023). We assume that both quiet and active regions have constant temperatures, which are set to  $T_{\text{eff}} = 5778\text{K}$  for the quiet photosphere and to  $T_{\text{eff}} = 6028\text{K}$  and  $T_{\text{eff}} = 5115\text{K}$  for the facula and spot regions, respectively (Meunier et al. 2010).

### 2.1. Active region properties

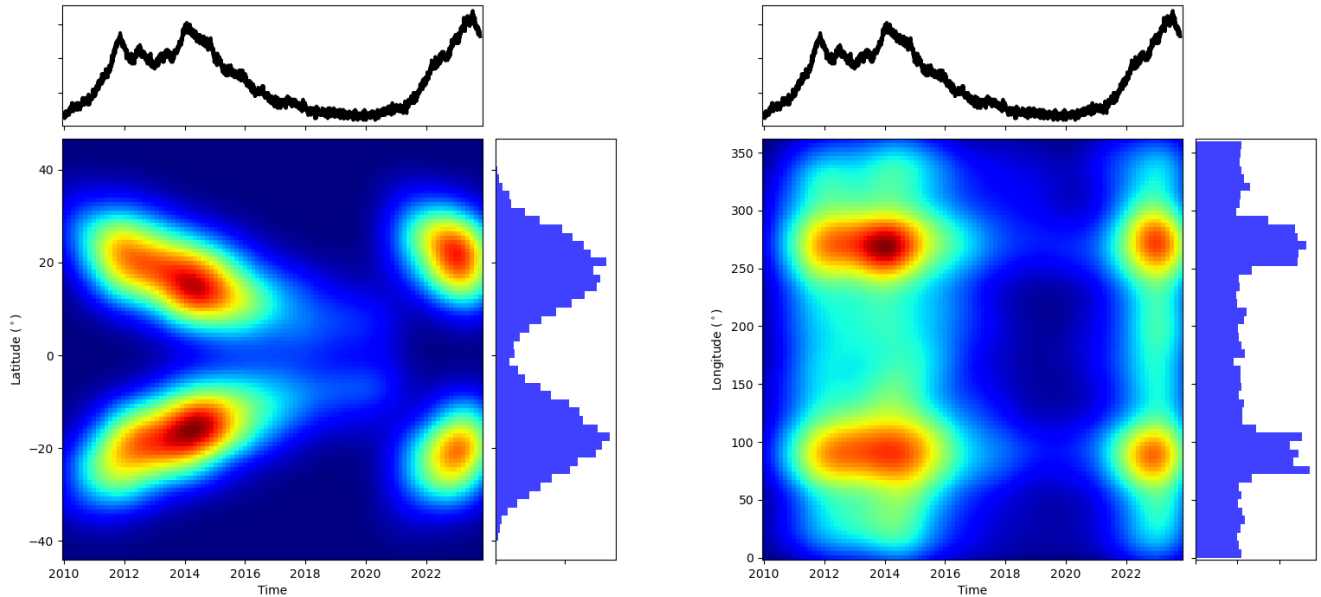
The number of spots on the solar surface is important in order to capture the short-term solar activity at the rotational and half of the solar rotational periods, and also the long-term activity induced by the solar magnetic cycle. For this study we used the spot number time series provided by the Solar Influences Data Analysis Center (SIDC) (SILSO World Data Center 2024).<sup>3</sup> The spot number  $n$  is described by the Wolf number  $R$ , which is defined by  $R = s + 10g$ , where  $s$  is the number of isolated spots and  $g$  is the number of complex spot groups. The numbers of these two spot types are obtained by using the fraction of isolated spots number versus all the spot numbers. In this simulation we used a value of 0.4, as mentioned in (Borgniet et al. 2015). So the numbers of isolated spots and complex spot groups are  $s = 0.4R$  and  $g = 0.06R$ , respectively. Another important aspect on the timescale of the solar magnetic cycle is the migration of active regions from high latitudes at the beginning of the cycle toward the equator at the end. We used the empirical equation in Hathaway (2011) to describe this butterfly diagram:

$$\bar{\lambda}(t) = 28^\circ \exp[-(t - t_0)/90]. \quad (1)$$

Here  $\bar{\lambda}$  is the average latitude of the active regions.  $t$  is time in units of months, and  $t_0$  is the cycle starting time. A magnetic cy-

<sup>2</sup> Data available at <https://doi.org/10.5281/zenodo.14262853> and code available at <https://github.com/YinanZhao21/SOAP-GPU>.

<sup>3</sup> Data available at <https://www.sidc.be/SILSO/home>.



**Fig. 1.** Location of active regions initialized with spot number evolution curve as a function of time. *Left:* Latitude evolution of injected active regions. The spot number evolution curve derived from the Solar Influences Analysis Data Center (SIDC) is shown in the top panel. The latitude distribution shown in the histogram on the right indicates that the north-south asymmetry factor is 0.5. The butterfly diagram is shown in the density map. *Right:* Longitude evolution of injected active regions. The spot number evolution curve derived from the SIDC is shown in the top panel. The longitude distribution shown in the histogram on the right indicates that the two considered active longitude regions are separated by 180 degrees.

cle of 132 months is used to repeat the butterfly diagram. A dispersion of  $6^\circ$  is added to the butterfly diagram to model realistic observations. As described in [Borgniet et al. \(2015\)](#), the active regions are not randomly distributed along the longitude of the solar disk. Two persistent active regions along the longitude are separated by  $180^\circ$ . The active longitudes are also included in our SOAP-GPU simulation. In [Borgniet et al. \(2015\)](#), the active longitude was formed by injecting a fraction of new sunspots into a confined area where existing sunspots were already present. For this paper we followed the same method. We defined two active longitude regions with average longitudes of  $90^\circ$  and  $270^\circ$  with a dispersion  $20^\circ$ . Forty percent of the spots are generated in the active regions, while the rest of the spots are uniformly generated along the longitude. The latitude distribution and longitude distribution of the active regions initialized by the 14 years of spot number data are illustrated in Fig. 1. Since we adopt the value of 11 years as the average duration of the magnetic cycle, only one butterfly diagram is fully shown in the latitude distribution. For simplicity, we chose to keep the active longitude regions fixed between cycles as an analysis by [Berdyugina & Usoskin \(2003\)](#) suggests that the same two active longitudes, separated by 180 degrees, have persisted on the Sun for over 120 years. While this may not hold for other stars, we also explored varying these longitudes between the two modeled cycles. As expected, the difference in the long-term is negligible as the inhibition of convective blueshift, related to the active region filling factor rather than its position on the disk, dominates the RV signal over magnetic cycles. The evolution of the individual active regions is also taken into account in the SOAP-GPU simulation. As [Borgniet et al. \(2015\)](#) pointed out, the timescale of the growing phase of active region is much smaller than the timescale of the solar rotation;

therefore, the growing phase is unlikely to contribute much to the solar activity on the timescale of the solar rotation. In the simulation, we thus only take into account the decay phase of active regions. The initial size of spot regions used in the simulation follows a log-normal distribution ([Baumann & Solanki 2005](#); [Borgniet et al. 2015](#); [Gilbertson et al. 2020](#)),

$$\frac{dN}{dA} = \frac{1}{\sigma_A \sqrt{2\pi}} \exp\left[-\frac{(\log(A) - \mu)^2}{2\sigma^2}\right], \quad (2)$$

where  $\sigma = \log(\sigma_s)$  and  $\mu = \log(\bar{A}_s)$ ,  $\bar{A}_s$  and  $\sigma_s$  are the mean initial size, and the corresponding standard deviation is in units of micro hemispheres ( $\mu\text{Hem}$ ). In the setup of the simulation we followed the same initialization as used in [Borgniet et al. \(2015\)](#): isolated spots have  $\bar{A}_s = 46.51$  and  $\sigma_s = 2.14$  and complex spot groups have  $\bar{A}_s = 90.24$  and  $\sigma_s = 2.49$ . We assumed that active regions follow a linear decay law. The average decay law also follows a log-normal distribution ([Martinez Pillet et al. 1993](#); [Borgniet et al. 2015](#); [Gilbertson et al. 2020](#)),

$$\frac{dN}{dD} = \frac{1}{\sigma_{\log D} D \sqrt{2\pi}} \exp\left[-\frac{(\log(D) - \mu_{\log D})^2}{2\sigma_{\log D}^2}\right], \quad (3)$$

where  $\sigma_{\log D} = \log(\sigma_D)$  and  $\mu_{\log D} = \log(\bar{D})$ ,  $\bar{D}$  and  $\sigma_D$  are the median decay rate, and the corresponding standard deviation is in units of  $\mu\text{Hem}/\text{day}$ . We used  $\bar{D} = 14.8, 30.9$  and  $\sigma_D = 2.01, 2.14$  for the isolated spots and complex spots group, respectively. For the simulation of facula regions, we assumed that every spot region is surrounded by a facula region. The initial size of the facula region is governed by the spot-to-facula ratio. We used a uniform distribution with a boundary from 5 to 10 to model the spot-to-facula ratio. The decay rate of faculae

also follows Eq. 3 with  $\bar{D} = 20.0$  and  $\sigma_D = 0.77$  (Borgniet et al. 2015). For every timestamp, the number of spots was obtained by using the corresponding spot decay rates. If the number of spots was smaller than the expected number derived from the SIDC data, we injected new spots. This method may lead to instances where the actual injected number exceeds the expected number. To quantify this effect we conducted a Monte Carlo simulation, and found that for spot numbers less than 20 there is a 20% excess injection, while for spot numbers greater than 20, the excess injection is only 2%. We agree that the relatively simple choices made in this section regarding active longitude properties, active region sizes, and their evolution can be discussed further in depth. However, the goal here is not to provide simulations tailored to match any specific star, but rather to demonstrate that with simple statistical assumptions it is possible to model realistic RVs for solar-type stars, including rotational and long-term magnetic effects, which is demonstrated in Fig. 3.

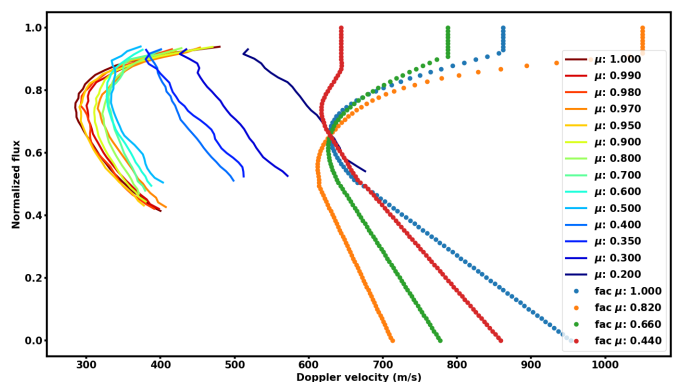
## 2.2. Input spectra properties

SOAP-GPU models, for each point on the visible solar hemisphere, the effect of stellar rotation, differential rotation, and limb-darkening. In addition, by injecting into SOAP-GPU different spectra as a function of  $\mu = \cos(\theta)$ , where  $\theta$  is the center-to-limb angle, it is also possible to model center-to-limb variations, as shown in Zhao & Dumusque (2023).

In practice, SOAP-GPU needs as input three spectral cubes for the quiet photosphere for faculae and for spots. As SOAP-GPU is able to model the Sun, and other stars as well, the code can use high-resolution stellar spectra from the PHOENIX database (Husser et al. 2013) as input. These spectra however do not include realistic convection and therefore spectral lines do not have a realistic shape, which is crucial to properly modeling spectra affected by stellar activity. As explained in more detail in Zhao & Dumusque (2023), we can use the quiet sun observations at different  $\mu$  angles from Löhner-Böttcher et al. (2019) to modify the bisector of PHOENIX spectra at different disk positions, and therefore obtain the desired spectral cube for quiet photospheric regions. Recently, Ellwarth et al. (2023) published the IAG solar atlas that includes high-resolution spectra of the quiet Sun taken at 14 heliocentric positions for the quiet solar disk, from  $\mu = 1$  (disk center) to  $\mu = 0.2$  (limb). It is therefore also possible to directly inject those spectra in the spectral cube for the quiet Sun. We differentiate between the two options, and refer to them as the PHOENIX and IAG cases.

For active regions, we used the results of Cavallini et al. (1985), who observed the bisector variation of the FeI line at 6301.5008Å as a function of  $\mu$ . This is the only observation available in the literature. Although using bisector changes from a single line may not perfectly model the CB effect since the convective blueshift inhibition may affect different lines in varying ways, this approximation should be valid to first order as all lines are redshifted and warped by the CB effect. As shown in Figs. 5 through 8 in Gray (2009), the selected lines in that paper generally follow common bisectors, with some departure due to line blending. We note, however, that the results from Gray (2009) should be interpreted with caution as the analysis by Palumbo et al. (2024) of observed spectra from the LARS spectrograph indicates that line bisectors of other lines (not presented in Gray 2009) differs quite significantly from the general trend, mainly for the very strong lines. Before injecting the appropriate bisectors as a function of  $\mu$ , the original bisectors in the input spectra are removed by fitting line bisectors with polynomial functions. A detailed description of bisector injection can be found in Sec-

tion 4.3.2 of Zhao & Dumusque (2023). In the case of a spot, a PHOENIX spectrum with  $T_{\text{eff}} = 5115\text{K}$  is used and, as in the case of the quiet photosphere, the bisector of spectral lines is changed to produce in the end the spectral cube for the spots. The spectral cube for the faculae is generated in the same way, just starting from a PHOENIX spectrum with  $T_{\text{eff}} = 6028\text{K}$ . We note that for both active regions, we have to adjust the velocity offset between quiet and active regions to properly take into account the inhibition of CB. This is done by assuming that the inhibition of convection in active region is the same as the maximum inhibition that happens in quiet regions between the center of the disk and the very limb (in our case at  $\mu = 0.2$ ). We can see in Fig. 2 the quiet bisectors derived from Ellwarth et al. (2023) and the active bisectors from Cavallini et al. (1985) shifted to match the maximum CB of the quiet spectra.<sup>4</sup>



**Fig. 2.** Line bisectors of quiet and active solar regions from the disk center ( $\mu = 1.0$ ) to the limb ( $\mu = 0.2$ ). *Continuous lines:* Line bisectors of quiet solar disk at 14 different heliocentric positions measured from Ellwarth et al. (2023), similar to Fig. 11 of Ellwarth et al. (2023). Unlike in Fig. 8 of Zhao & Dumusque (2023), we do not inject the bisectors of the quiet region into the spectra, as they are naturally included. *Dotted lines:* Fit of the bisectors of the FeI 6301.5008Å spectral line inside a facula region, as measured by the Fabry-Perot interferometer at the Donati Solar Tower (Cavallini et al. 1985). Below a depth of 0.5, a linear fit is performed, while a fifth-order polynomial is used to model the top part of the bisector. To prevent unrealistic values when interpolating the polynomial above a normalized flux of 0.9 where no measurement exists, we selected the most redshifted part of the top bisector, explaining the vertical values for very shallow depths. The bisector of active regions at different  $\mu$  angles are all shifted by 350 m/s based on our hypothesis that convection is fully suppressed in magnetic regions.

## 2.3. Spectral time series results

Using SOAP-GPU with the inputs described above, we simulated one solar spectrum every two days for 14 years, between 2010 and 2024, to cover solar magnetic cycle number 24 and the HARPS-N solar observations. Based on an active region map that is generated at each timestamp from the spot number, SOAP-GPU integrates over all active regions at the given time to generate an integrated spectrum, which corresponds to the difference between the quiet and active Sun at the position of active regions. This integrated spectrum is then subtracted from the integrated spectrum of the full disk covered with quiet regions only, therefore producing the full disk integrated spectrum

<sup>4</sup> When the quiet Sun spectral cube is derived in the PHOENIX case, the quiet Sun bisectors are slightly different from the ones derived from the IAG case, and therefore the CB inhibition of the active region bisectors will be slightly different.

affected by activity. In total, our simulation represents  $2555 \times 3$  spectra as SOAP-GPU gives as output the solar spectrum affected by all activity components, but also affected either by the inhibition of CB or by the flux effect. The RV values are derived by cross-correlating the obtained spectra with the ESPRESSO G2 mask. The RV time series of different stellar activity effects and the corresponding periodograms are shown in Fig. 3.

Based on the simulation using only the spot number, the stellar activity affected RVs are dominated by the inhibition of CB, as presented in Fig. 3. Overall, the difference between the PHOENIX and IAG cases is very minor. This is not surprising since both simulations used the same set of active region maps generated from the spot number. Since the simulation covers the HARPS-N solar observation, we also compared the simulated RVs with the HARPS-N solar RVs. Although the real location and size of active regions for a given epoch cannot be modeled by only using the spot number and empirical equations, the long-term magnetic cycle presented in the HARPS-N solar data can be well modeled by the simulations, as shown in the top left and top right panels of Fig. 3. The periodograms of the total effect and the CB effect are dominated by a long-term trend associated with the magnetic cycle, implying that in the long-term the CB effect dominates over the flux effect. The periodogram of the flux effect is itself dominated by stellar activity at the rotational timescale and its respective harmonics (one-half and one-third of the solar rotation). Once the long-term trend is removed by fitting a third-order polynomial on the RVs of the total and CB effects, stellar activity at the rotational period and the respective harmonics stands out, as in the flux effect case.

### 3. Solar spectra modeling using SDO data

As SOAP-GPU requires as input an active region map at each timestamp (see Sect. 2.3), we can inject for each time of the simulation the observed location and size of the solar active regions (spots and faculae) as extracted from the different images produced by SDO. Compared with the simulation that uses only the spot number time series as input (see Sect. 2), our simulation using SDO image information as input does not require additional assumptions on the initial conditions of the active regions, active region evolution curves, and parameterization of latitude and longitude distributions since all those information are included in the SDO images. By using SDO data, SOAP-GPU is able to model the activity-affected spectra more precisely than the simulation generated with only the spot number time series, which allows us to then compare in detail the simulated spectral time series with the HARPS-N solar spectra. In Sect 3.1, we describe the different processing steps of SDO images to derive the active region map used as input to SOAP-GPU. The results of the simulation are present in Sect 3.2. We compare the SDO simulation with HARPS-N solar spectra at the spectral level in Sect 3.3.

#### 3.1. SDO data preprocessing

Haywood et al. (2016) extracted the solar active region properties using the dopplergram, intensity continuum, and magnetogram images from the Helioseismic and Magnetic Imager (HMI, Schou et al. 2012) on board SDO. We followed a similar approach to generate the active region map as input of SOAP-GPU. Since limb darkening is already included in SOAP-GPU, we only used the flattened intensitygram data to localize the spot regions and used the magnetogram map to localize the facula regions, as shown in the top panels of Fig. 4. The original SDO

FITS files were downloaded from the Joint Science Operations Center (JSOC).<sup>5</sup> We used the magnetogram and flattened intensitygram data with a cadence of one day. We downloaded the data from January 1, 2015, to August 15, 2024, which covers the current HARPS-N solar observations (July 2015 to February 2024). The HARPS-N solar data that we used here are binned per day, and therefore we have one frame per day, for a total of 3494 active region maps.

Given that the original SDO data have a dimension of 3900 by 3900, it is computationally expensive for SOAP-GPU to use such high-resolution maps as input to perform the simulation. We therefore lowered the resolution of the flattened intensitygram and magnetogram from 3900 by 3900 to 300 by 300 for computational efficiency. We then detected spots as regions in the flattened intensitygram for which the contrast is less than 0.89 (Haywood et al. 2016). Regarding the detection of faculae, we adopted for the magnetogram the same threshold value of  $24.0/\mu$  as mentioned in Haywood et al. (2016). Pixels above this value were considered facula regions. The derived high-resolution active region map is illustrated in the bottom left panel of Fig. 4.

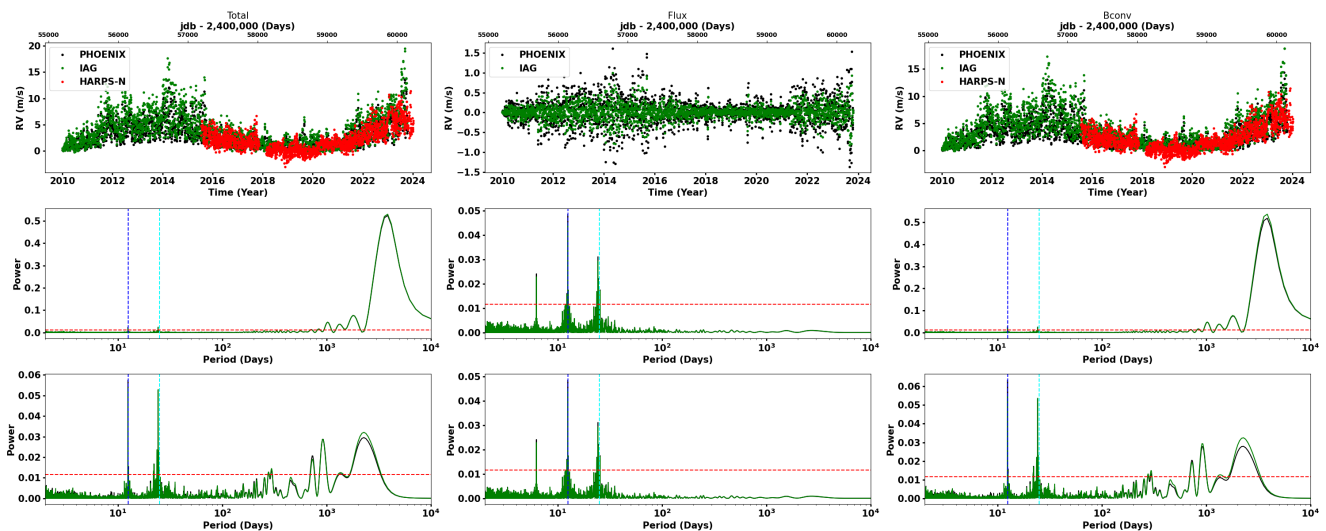
We lowered the resolution of the derived active region map from 3900 by 3900 to 300 by 300, which means that a subgrid of 13 by 13 in the high-resolution map is converted to 1 pixel in the low-resolution map. If more than 50% of the low-resolution pixels were covered by a facula region in the high-resolution map, then the low-resolution pixel was considered to be covered by a facule. We repeated the same process to lower the resolution of the flat intensitygram. Once we derived both the low-resolution facula and spot maps at each timestamp, we combined the two into one active region map. If there were overlapping pixels, we counted them as spot pixels since spot regions are surrounded by facula regions. An example of a final low-resolution active region map used as input to SOAP-GPU is illustrated in the bottom right of Fig. 4.

We note that lowering the resolution of the SDO maps could have an impact on the final modeled RVs (see Sect. 3.2) if active regions start to be truncated. However, we compared the derived RVs with the full and reduce resolution map of SDO and the results were extremely similar, demonstrating that the resolution that we chose was appropriate.

#### 3.2. Spectral time series results

We ran the SDO image-based simulation with the two sets of the input seed spectral cubes: the PHOENIX and the IAG cases. For each SDO active region map, there are three spectra simulated to account for the flux effect, inhibition of CB effect, and the total effect. The corresponding RV data points were obtained by using the ESPRESSO G2 mask. Compared with the simulation using only spot numbers, the simulation using the SDO maps contains accurate information about the active regions, which allows us to directly compare the results with the solar observation. We used the solar spectra observed by the HARPS-N solar telescope ranging from 2015 July 19 to 2024 January 8 (Dumusque et al. 2021, Dumusque et al. in prep.). We collected 2040 daily binned solar spectra. All the solar spectra are preprocessed by the YARARA data pipeline (Cretignier et al. 2021) to remove systemic errors such as those induced by tellurics, ghosts, interference patterns, and stellar activity. Since we need to compare the simulated stellar activity with the observed one, we injected the YARARA stellar activity correction component back into the YARARA-

<sup>5</sup> <http://jsoc.stanford.edu/ajax/lookdata.html>



**Fig. 3.** RV time series and periodograms from the solar simulations using spot number as input. In each panel we show the results for the PHOENIX and IAG cases in black and green, respectively. *Top:* From left to right, we show the RVs of the total effect, flux effect, and the effect of CB inhibition. The RVs of the HARPS-N solar observation, in red, are also illustrated in the left and right panels for comparison. *Middle:* Corresponding periodograms for each effect. The FAP level of 0.1% is represented by the red dashed line. The half rotation and the rotation period of the Sun are indicated with blue and cyan dashed lines. *Bottom:* Corresponding periodograms of each effect, but with the magnetic cycle removed. Since the RVs from the flux effect are not affected by the long-term trend induced by the magnetic cycle, the periodogram remains the same.

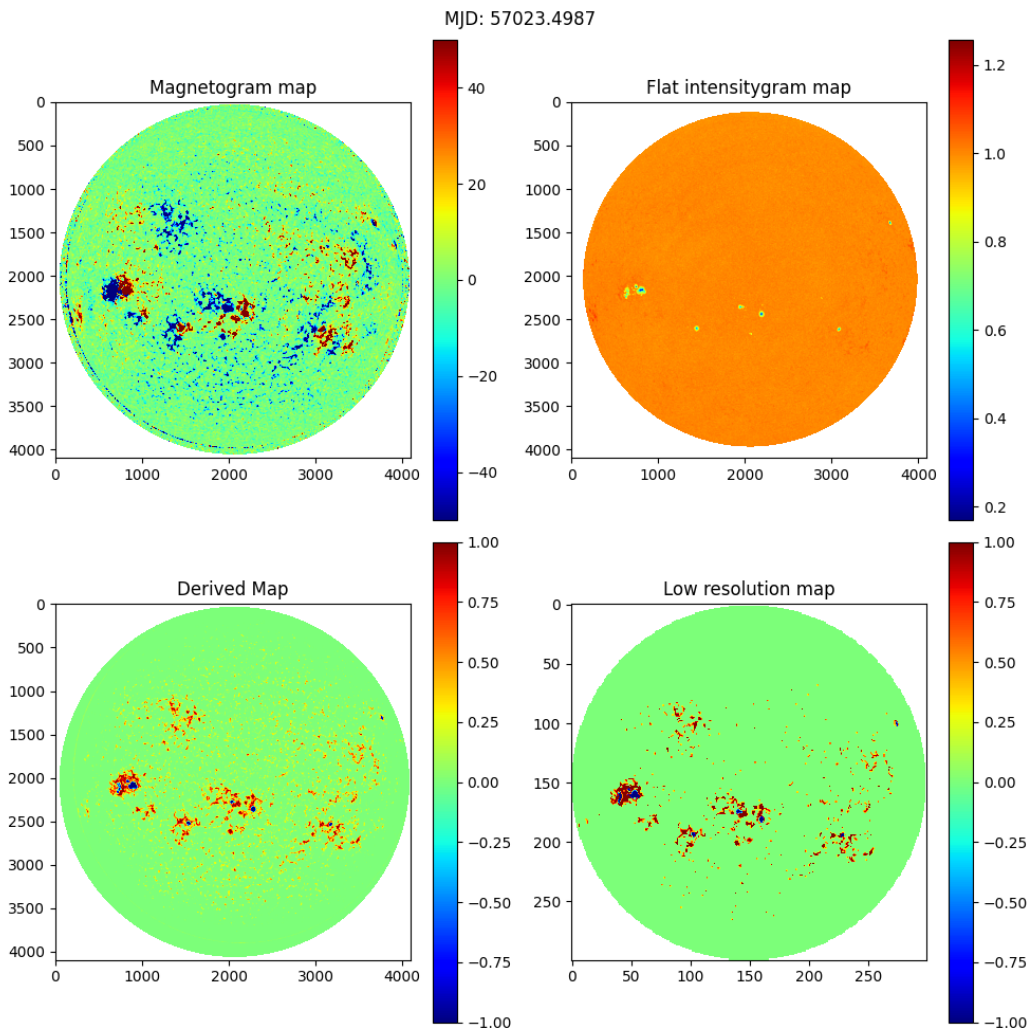
cleaned spectra. We derived the RV time series of HARPS-N solar data by using the same ESPRESSO G2 mask as the simulated spectra. Since the SDO images and the HARPS-N solar observation have different sampling rates, we interpolate the simulated RV time series into the observation grid of HARPS-N solar data. The resulted RV time series for the PHEONIX and IAG cases are illustrated in Figs. 5 and A.1, respectively.

We divide the RV time series into three sections: the end of cycle 24 decreasing activity phase, the quiet phase, and the new cycle 25 increasing activity phase. By comparing the simulated RVs with the observed ones and deriving the RV rms by fitting the offset for each phase, we found for the decreasing and increasing activity phases that the RV rms can be strongly reduced from 1.27 to 0.89 and 1.99 to 0.78 m/s, respectively. For the quiet phase the RV rms is only marginally reduced from 1.09 to 0.91 m/s which is not surprising as the SDO model is rather flat during this time due to the weak solar activity. This demonstrates that our SOAP-GPU SDO simulation is able to model well the observed RVs as stellar activity is mitigated to below 1 m/s level during the active phases. Regarding the increasing activity phase, we see that the residuals after correction of the SDO model are at the level of 0.78 m/s, which is comparable to what is predicted from supergranulation (e.g., 0.86 and 0.68 m/s in Lakeland et al. 2024; Al Moulla et al. 2023, respectively). For the decreasing activity phase, the RV rms residual is 0.89 m/s, which is slightly higher than what is predicted from other stellar signals and the residuals clearly show a trend that is likely responsible for this extra jitter (removing this trend lowers the RV rms to 0.84 m/s). We do not know the origin of this trend; however, it could be due to a different convective blueshift from one cycle to the next, or a different strength in its inhibition, which is a constant value equal to 315 m/s in our model. Finally, the RV rms of the residuals during the quiet phase is also greater than that predicted from supergranulation, suggesting that instrumental systematics are clearly at play. Two rather strong deviations compared to the mean RV around BJD 2458500 and 2459000 in the quiet phase are caused by a HARPS-N detector warm-up and by the change

of the main ThAr lamp used for wavelength solution, respectively. The first intervention slightly changes the point spread function on the detector and induces a small RV shift; the second slightly changes the thorium spectrum, therefore affecting the wavelength solution and thus also inducing a RV offset. In addition, the gap in the data around BJD=2458000 corresponds to the fiber injecting light from the solar telescope to the calibration unit being damaged. After replacing it, we observed an RV offset of about 0.5 to 1 m/s. As the problem happened before injection into the calibration unit, it is difficult to understand how a different injection at this level could introduce an RV offset as in principle the octagonal fibers used by HARPS-N, in addition to a double scrambler, should strongly mitigate any difference in light-injection at the calibration unit level. However, a detector warm-up also happened within this observational gap and it could be the cause of this RV offset. When investigating the behavior of the RV around each warm-up, it seems that not all of them have a significant impact on the RVs. A detailed study on the effect of warm-ups is beyond the scope of this paper, but should be investigated in the future to understand precisely what is happening and hopefully find a way to mitigate the impact of warm-ups on RV precision.

### 3.3. Comparison with HARPS-N data at the spectral level

The good agreement between the simulated spectra and the HARPS-N spectra in the RV space inspires us to further compare the two datasets at the spectral level. Given that the CCF is derived from a specific line selection, in our case from the ESPRESSO G2 mask, only those lines contribute to the corresponding RV value. Because the RV signal is measured from the spectral lines and not the stellar continuum, we decided to investigate the morphology of spectral lines in the two datasets rather than to compare the full spectra. Here we used the line list derived from Cretignier et al. (2020). Blended lines and weak lines were excluded from the analysis due to contamination and noise. We only used lines that are deeper than 0.1 in normalized



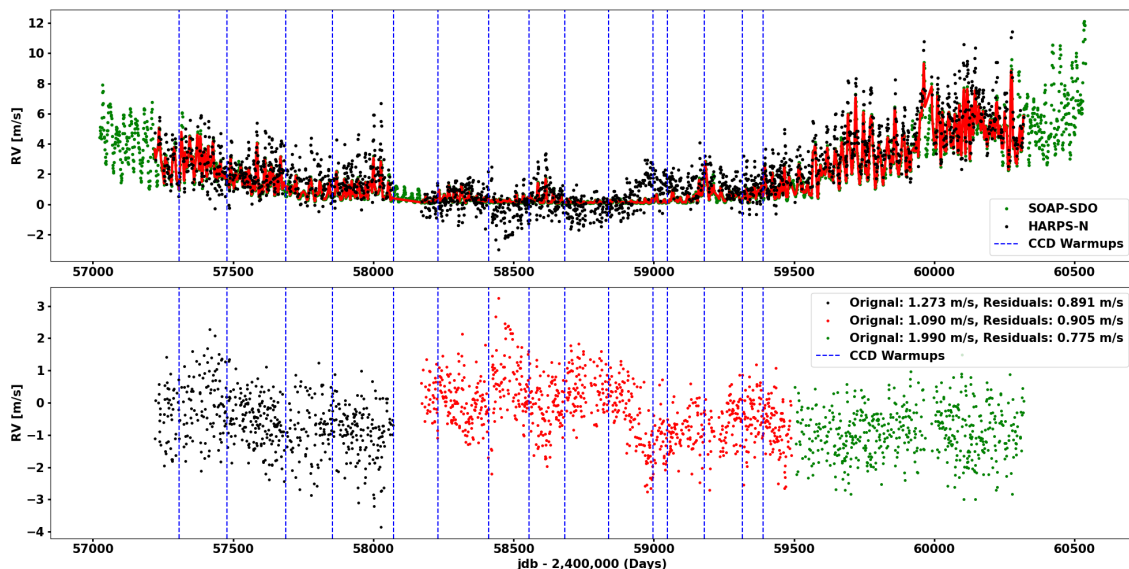
**Fig. 4.** Pre-processing of SDO magnetogram and flattened intensitygram images to derive active regions. Top left: Original SDO magnetogram image. Regions with positive and negative magnetic fields are labeled in red and blue, respectively. Top right: Original SDO flattened intensitygram image. Regions with high contrast (in green) indicate the presence of spots. Bottom left: Derived high-resolution active region location with a dimension of  $3900 \times 3900$ . The facula regions are in red, while the spots are in blue. Bottom right: Low-resolution map of active regions with a size of  $300 \times 300$ , used as input of SOAP-GPU to speed up computation.

flux. For each selected spectral line, we compared in the model and in the real observations how the line Doppler shift and depth behave over time.

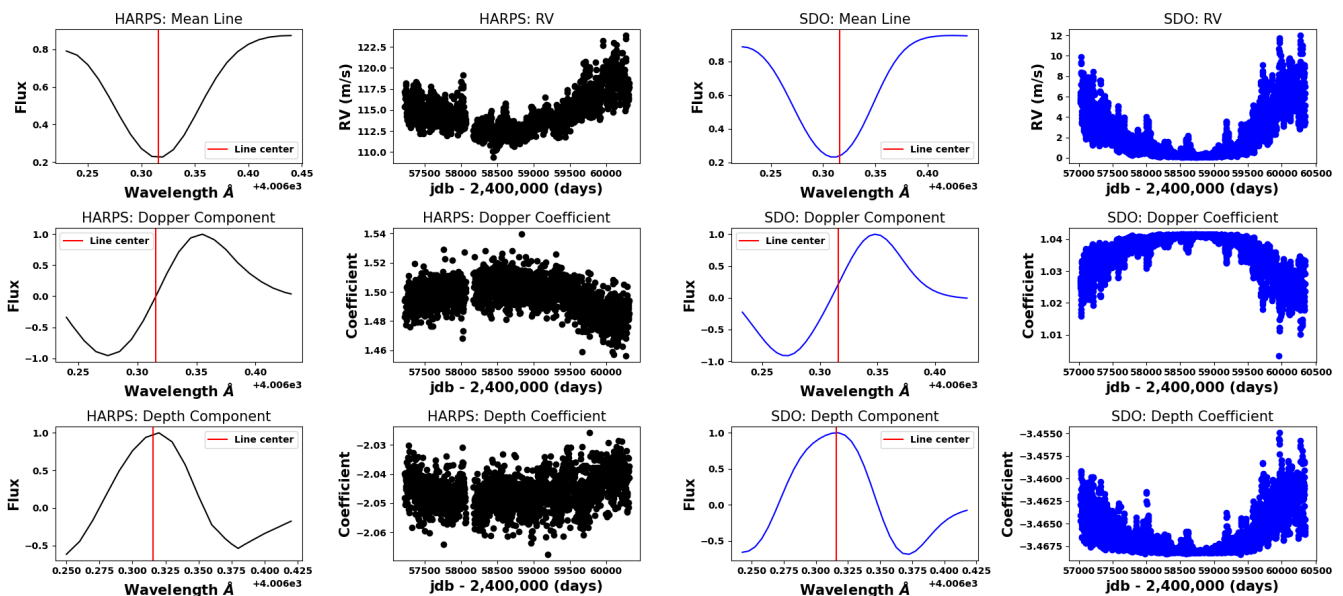
To investigate the line profile in the line Doppler shift and line shape domains, we first extracted the line shift and line depth basis that was then used to derive the coefficient of the line profile projected onto this basis. As the Doppler shift is proportional to the first derivative (e.g., Bouchy et al. 2001), we used the first-order derivative of the mean line profile as the Doppler shift basis. Regarding the line-depth basis, we can show that the second-order derivative is sensitive to line depth changes. Therefore, we used the second-order derivative of the mean line profile as the line-depth basis. In the first and the third columns of Fig. 6, we show in the case of one spectral line the mean line profile as well as the line shift basis and line depth basis.

In order to quantify the line shift and line depth changes for a given spectral line as a function of time, we projected the line profile time series onto the line shift and line depth basis and looked at the obtained respective coefficients. We show the coefficient time series for the FeI line at  $4006.31 \text{ \AA}$  in the second and fourth columns of Fig. 6. In both the HARPS-N and

the SDO-based SOAP-GPU simulated spectral time series, the line shift coefficients and the line depth coefficients time series show a strong correlation with the corresponding RV time series. This is not surprising since stellar activity modifies the line shape, which can be decomposed as a shift and depth variation. To measure the similarity in the variation of line morphology between the HARPS-N spectra and the SDO-based SOAP-GPU spectra, we first computed for each spectral line and for the real and simulated cases the time series of the coefficients obtained when projecting onto the line shift or the depth basis. We then computed for each spectral line the Pearson correlation coefficient between the obtained coefficient time series. A line with a large positive Pearson correlation coefficient value indicates that its shift and/or depth variations are well modeled by SOAP-GPU. Some lines, however, present large negative Pearson correlation coefficient values, for example the line at  $4512.74 \text{ \AA}$  shown for both the IAG case and PHOENIX case in the third row of Fig. 7 and Fig. 8, respectively. The Pearson correlation coefficient value of the depth coefficient for the IAG case shows a strong anti-correlation of  $-0.593$ , while the correlation is  $0.652$  for the PHOENIX case. Given that the inhibition of convective



**Fig. 5.** RV data of simulated spectral time series modeled with SDO images. The input seed spectra generated from the PHOENIX synthetic spectral library are used as the inputs of SOAP-GPU (see Sect. 2.2). *Top:* RV comparison between the SOAP-GPU simulated spectra and HARPS-N solar spectra. The RVs derived from the simulated spectra are highlighted in green while the RVs derived from HARPS-N solar spectra are shown in black. We interpolated the simulated RV time series (in red) to match the HARPS-N solar observation. The blue dashed vertical lines indicate the times of the CCD warm-ups. *Bottom:* Residuals after subtracting simulated RVs from the HARPS-N solar RVs. The RV residuals are divided into three phases: the decreasing cycle 24 activity phase (black), the quiet phase (red), and the new increasing activity phase of cycle 25 (green). The corresponding rms are 0.891 m/s, 0.905 m/s, and 0.775 m/s, respectively. The rms of HARPS-N solar spectra for these three phases are 1.273 m/s, 1.090 m/s, and 1.990 m/s, and are shown here for comparison.

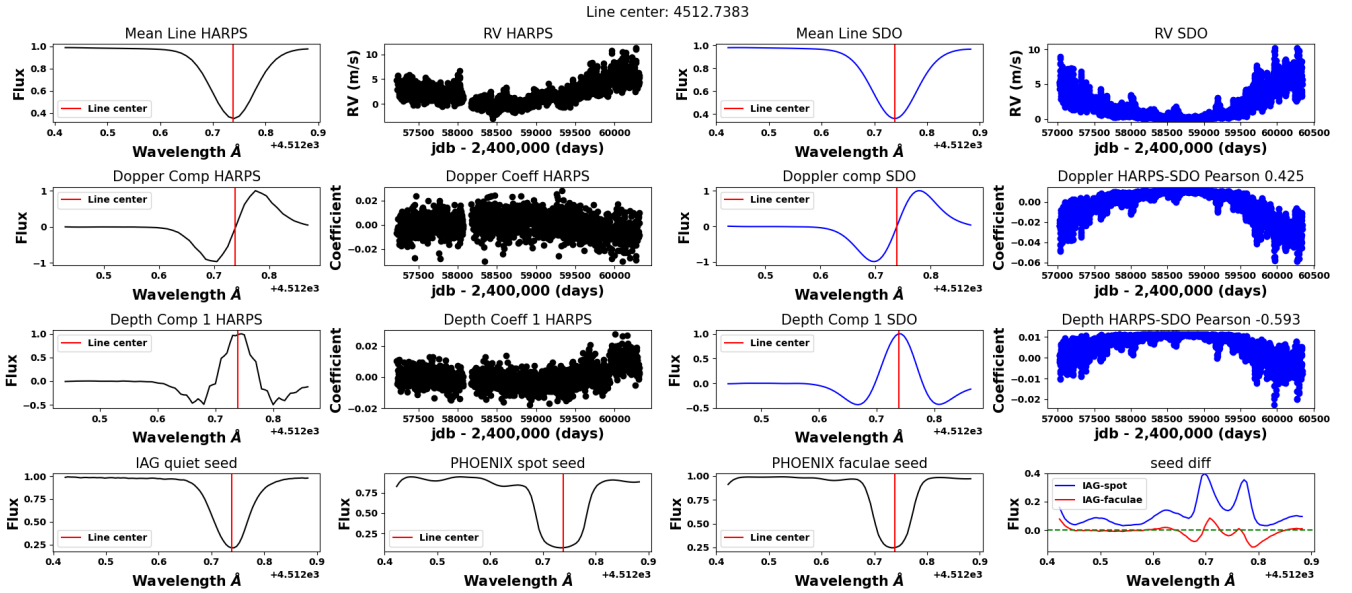


**Fig. 6.** Example of line profile decomposition in the Doppler-shift and line-depth domains for the FeI line at  $4006.31 \text{ \AA}$ . *First column:* Line profiles of HARPS-N solar spectra. The top panel is the average line profile. The middle panel is the Doppler component of that line. The bottom panel is the depth component of the line. *Second column:* Top panel is the RVs measured from the CCFs. The coefficients associated with the Doppler and depth components are shown in the middle and bottom panels, respectively. *Third and fourth columns:* Same as the first and second columns but for the simulated spectral line. The Pearson correlation between the line in HARPS-N solar spectra and the same line in the simulated spectra is 0.70 for the Doppler component and 0.41 for the depth component

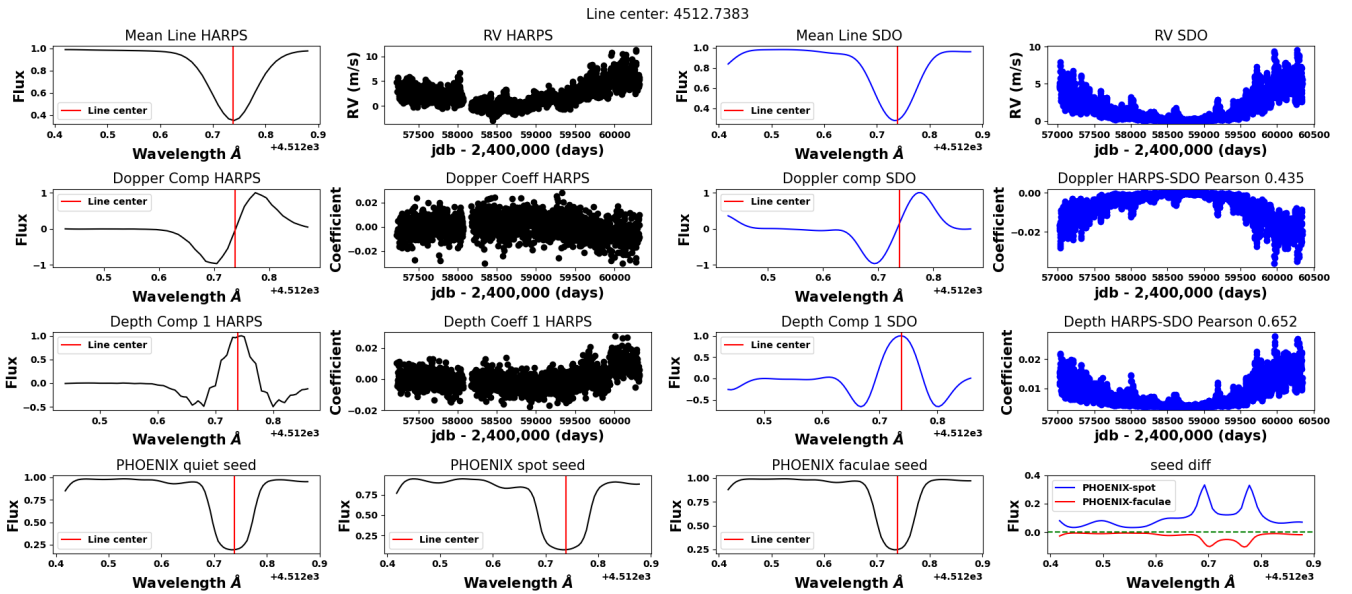
blueshift is the main contribution to solar activity (e.g., Meunier et al. 2010; Liebing et al. 2021), this anti-correlation is likely due to the spectral difference between the input seed spectrum of the quiet Sun and the input seed spectrum of active regions, as described by Equation 2 and Equation 10 in Zhao & Dumusque (2023). From Fig. 4 we know that facula regions are much larger than spot regions, making faculae the main contributor to the

convective blueshift inhibition. The spectral difference between the input seed spectrum of the quiet Sun at  $\mu = 1.0$  and the input seed spectrum of a facula (hereafter  $\Delta S_{bcnv,quiet-faculae}$ ) is illustrated in the last panels of the fourth row in Fig. 7 and Fig. 8. Faculae are hotter than the quiet Sun, and therefore we expect the spectral lines to be shallower due to the higher temperature. The value of  $\Delta S_{bcnv,quiet-faculae}$  should therefore be smaller than





**Fig. 7.** Similar to Fig. 6, but for the Fe I line at  $4512.74 \text{ \AA}$  from the IAG case. It has a negative Pearson correlation coefficient value of  $-0.593$  for the depth coefficients. An additional row is added to highlight the input seed spectral line profile used in this simulation. *Fourth row:* Line profile of the IAG spectrum at  $\mu = 1.0$  for modeling the quiet Sun and line profiles from the PHOENIX spectral library for modeling spot and facula regions. The last panel in this row shows the line profile difference between the input seed spectrum of the quiet Sun and the input seed spectra of the active regions. A horizontal green dashed line indicates the zero level. The profile of  $\Delta S_{bconv,quiet-faculae}$  has regions greater than zero, which is not expected.

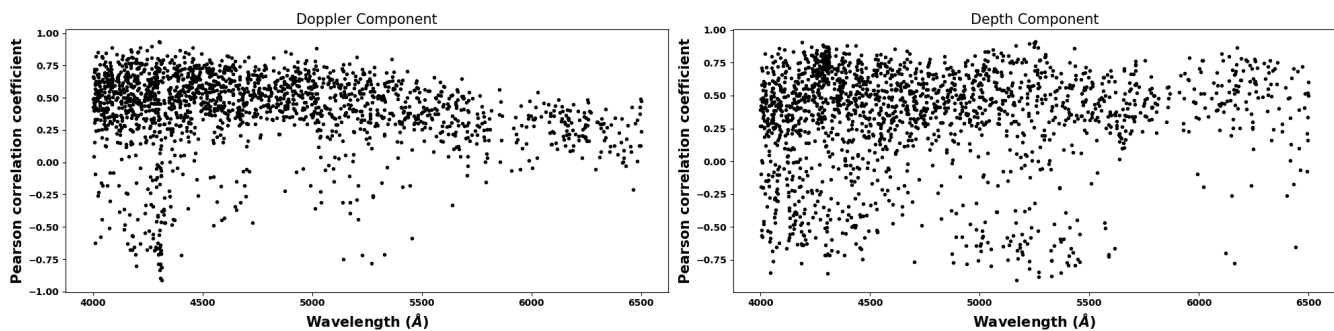


**Fig. 8.** Same as Fig. 7, for the same line, but for the PHOENIX case. The correlation of depth coefficient shows a positive Pearson correlation coefficient value of  $0.652$  here. The last panel in the fourth row shows the difference between the input seed spectrum of quiet region. The profile of  $\Delta S_{bconv,quiet-faculae}$  is always below zero as expected.

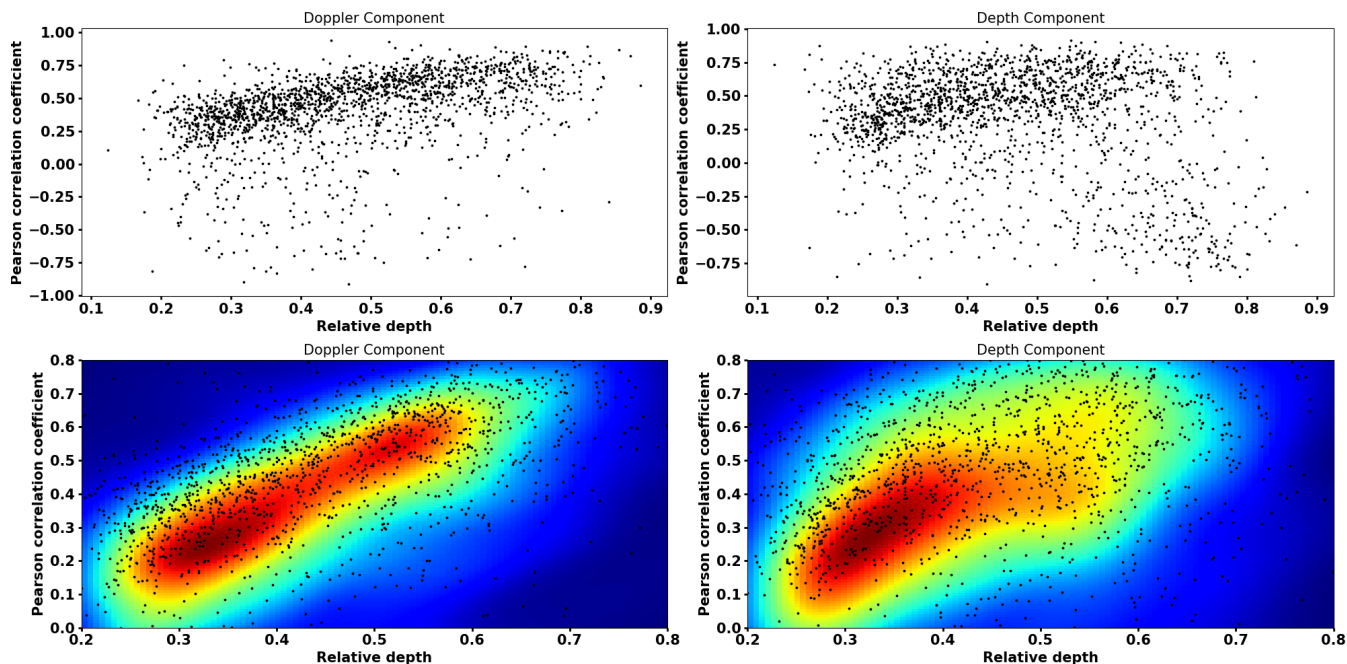
zero. This is the case for the SOAP-GPU simulation using as input for the quiet Sun the PHOENIX synthetic library, as shown in Fig. 8. However, for the IAG case in Fig. 7,  $\Delta S_{bconv,quiet-faculae}$  for the line at  $4512.74 \text{ \AA}$  goes above zero near the line core. Given that the temperature difference between the quiet region and the facula region is smaller than  $250\text{K}$  ( $250\text{K}$  at the limb,  $33\text{K}$  at disk center Meunier et al. 2010), there is an inconsistency between the IAG observations and PHOENIX synthetic library for this small temperature difference. In other words, a mismatch between the line gradient and the effective temperature gradient

for certain lines will lead to an anti-correlation between the real and simulated cases. It is therefore dangerous to mix input spectra seeds coming from real observations and synthetic libraries, and for that reason, as long as we do not have good observations for spots and faculae, we do not recommend using the IAG input seed spectra with SOAP-GPU.

We investigated the line similarity in different physical domains. We first investigated the line similarity in the wavelength domain, as shown in Fig. 9 and Fig. A.2. We find that there is a linear trend between the Pearson correlation coefficient value



**Fig. 9.** Line similarity in the wavelength domain for the SDO-based SOAP-GPU simulation using the PHOENIX input seed spectra (see Sect. 2.2). *Left:* Similarity of line profiles associated with the Doppler component. Line similarity decreases with wavelength, and therefore it seems that the SDO-based SOAP-GPU simulation can model blue lines better than red lines. This is likely due to telluric line contamination at the red part of the HARPS-N solar spectra. *Right:* Similarity of line profiles associated with the depth component. There is no correlation between the Pearson correlation coefficient value and wavelength.



**Fig. 10.** Line similarity in the relative line depth domain for the SDO-based SOAP-GPU simulation using the PHOENIX input seed spectra (see Sect. 2.2). *Top left:* Similarity of line profiles associated with the Doppler component. *Bottom left:* Similar to the top left panel, but displaying only lines with Pearson correlation coefficient values greater than 0.0. Kernel density estimation is used for better visualization of the data distribution. There is a strong linear correlation between the line profile similarity and the relative line depth as the strong lines show higher similarity than the weak lines. This is likely because the strong lines have a higher signal-to-noise ratio than the weak lines. *Top right:* Similarity of line profiles associated with the depth component. *Bottom right:* Similar to the top right panel, but displaying only lines with Pearson correlation coefficient values greater than 0.0. Kernel density estimation is used for better visualization of the data distribution. A strong linear correlation is visible at relative line depth smaller than 0.4, where the lines are photon-limited. Beyond 0.4, the data is more uniformly distributed.

and the wavelength for the line shift component. As the wavelength increases, the Pearson correlation coefficient value decreases, as shown in the left panels of Fig. 9 and Fig. A.2. This means that the difference between the HARPS-N solar spectra and the SDO-based SOAP-GPU simulated spectra is larger in the red part of the spectrum than in the blue part. One possible explanation is the presence of telluric features. Since telluric lines start to pollute the observed HARPS-N spectra at a wavelength of  $\sim 5000$  Å, the redder absorption lines suffer more from telluric line contamination than bluer spectral lines. In the line depth domain, the Pearson correlation coefficient values are uniformly distributed in the wavelength space and we cannot see any linear trend. One possible reason is that the contamination

from telluric line affects the line shift more than the line depth as telluric lines contaminating any portion of a stellar spectral line will induce a Doppler shift, while only contamination at the precise location of the line core will affect line depth. As described in the previous paragraph, lines with negative Pearson correlation coefficient values indicate a mismatch between the line gradient and the effective temperature gradient.

We further investigated the line similarity in the line relative depth domain. The Pearson correlation coefficient value shows a linear trend with the line relative depth for the Doppler component. As the relative depth increases, the line similarity between the HARPS-N solar spectra and the SDO simulated spectra also increases for the line shift variation with only a small dispersion

to negative Pearson correlation coefficient values, as shown in the left panels of Fig. 10. For the IAG case, there is a strong correlation for lines with positive Pearson correlation coefficient values, as shown in the bottom left panel of Fig. A.3. Given the inconsistency between the IAG observation and PHOENIX synthetic library described above, the number of lines with negative Pearson correlation coefficient values in the IAG case is much larger than in the PHOENIX case. The possible explanation for this linear trend is that the morphology measurement on weak lines in the HARPS-N solar spectra is greatly affected by photon noise and the signal-to-noise ratio (S/N) of those weak lines is not sufficient to derive precise measurements. This is something that was also observed and explored in depth in the study of Cretignier et al. (2020).

For the line similarity of the depth component, there is also a strong correlation between the Pearson correlation coefficient value and the corresponding relative depth: only a small fraction of lines have negative Pearson correlation coefficient values, as shown in the top right panel of Fig. 10. However, when lines with negative values are excluded, we found that the correlation is not linear across the entire relative depth range when the input seed spectra for SOAP-GPU are based on the PHOENIX library, as shown in the bottom right panel of Fig. 10. For lines with a relative depth less than 0.4, there is a strong correlation between the relative depth and the line morphology, which is likely due, as in the case of the line shift, to the low S/N of weak lines in the HARPS-N solar spectra. For lines with relative depth greater than 0.4, in the PHOENIX case (see Fig. 10), the slope of this linear trend becomes smaller and the data are more widely distributed. When the input seed spectra for SOAP-GPU are based on the IAG solar atlas, we found that both lines with positive and negative Pearson correlation coefficient values show a strong correlation (see the top right of Fig. A.3), similar to the case for the wavelength domain. After excluding lines with negative Pearson correlation coefficient values, we observed a strong correlation with line depth for all lines, again pointing, as in the case of the shift analysis, to a S/N dependence, as shown in the bottom right of Fig. A.3. The significant difference observed between the PHOENIX and the IAG simulated spectral time series for lines with positive Pearson correlation coefficient values, likely points toward a modeling of the depth of stellar lines that is not fully correct in the PHOENIX case. As explained in Sect. 2.2, PHOENIX spectra do not include the proper convective blueshift as is naturally the case for the IAG solar atlas. Therefore, we artificially added convective blueshift by modifying the line bisector based on a few solar line measurements (Löhner-Böttcher et al. 2019). We therefore assume that all lines have the same bisector, which is probably correct to first order for unblended lines, but will not work for blended lines. Although we make a careful selection of lines to reject evident blends, to a certain extent all lines are blended, and therefore having an impact on the way we model the line bisector.

#### 4. Discussion and conclusion

In this paper we demonstrated two methods of using SOAP-GPU to simulate realistic spectral time series affected by stellar activity perturbations for the Sun, and publish the simulations for the benefit of the community working on stellar activity mitigation in RV measurements. SOAP-GPU allows users to simulate solar spectral time series using as input the number of spots as the function of time. Other physical processes, for example the active size evolution curve, the spatial migration, and the initial size distribution, are parameterized by empirical equations de-

rived from previous solar observations. The input spectral cube used to model the quiet solar surface as a function of position on the limb ( $\mu$  angle) can either come from PHOENIX synthetic spectra on which the bisector of spectral line is modified as a function of  $\mu$ , based on solar observations from Löhner-Böttcher et al. (2019), or from observations of the full visible spectrum taken at different heliocentric positions (Ellwarth et al. 2023).

By using the spot number as a function of time to model the spectral time series affected by solar activity, the prominent features of solar activity can be well modeled. The long-term magnetic cycle and the signals associated with solar rotation can clearly be observed in the derived RVs from the simulated spectra (see Fig. 3). We are therefore confident that such a strategy can be used to realistically model stellar activity on stars other than the Sun. However, when going far from the Sun in spectral type, although the use of PHOENIX spectra and the modification of the convective blueshift can adapt to different spectral types (see Zhao & Dumusque 2023), we should be careful that the active region properties (initial size, decay rate, latitude migration, and the effective temperatures of active regions) are based on solar observations and may differ for other spectral types.

In order to precisely model solar activity and compare with the corresponding HARPS-N solar spectra, we can also give as input to SOAP-GPU the magnetograms and flat intensity maps from SDO observations. Unlike generating the location of active regions with the spot number curve and empirical equations from solar observations, the active region location and size at different timestamps can be precisely extracted, which allows us to compare on a one-to-one basis the simulated spectra with the HARPS-N solar observations.

In the RV space, the time series of SDO simulated spectra is in good agreement with the RVs derived from the HARPS-N solar spectra, with a residual RV rms of  $\sim 0.8 - 0.9$  m/s. In our analysis, we split the time series into three parts, two corresponding to active phases, and one to the inactive phase. We clearly demonstrate that our SOAP-GPU SDO model is able to significantly mitigate stellar activity during the decreasing and increasing active phases to levels below the level of 1 m/s, from 1.27 and 1.99 m/s down to 0.89 and 0.78 m/s, respectively. While the increasing active phase show a RV rms of 0.78 m/s similar to what is expected from supergranulation, the decreasing phase still show higher jitter, 0.89 m/s, likely due to a trend in the residuals, of unknown origin. We also see some significant residuals during the inactive phase (RV rms = 0.91 m/s), which is likely due to residual instrumental systematics. A detailed study of those systematics is beyond the scope of this paper, but the initial investigation points toward warm-ups of the detector and change of the master thorium-argon lamp used for the wavelength solution.

At the spectral level, we compared spectral lines in the line shift and line depth domain. We measured the similarity in line morphology variation between the two spectral datasets by projecting each line onto a shift and depth basis and comparing the Pearson correlation between the obtained coefficients. Some of the lines show negative Pearson correlation coefficient values, indicating a mismatch between the line gradient and the effective temperature gradient. Although the necessary physics are naturally included in the IAG observations, the number of lines with negative Pearson correlation coefficient values in the IAG case is larger than in the PHOENIX case. This is due to an inconsistency between the IAG observations and the PHOENIX synthetic library which is always used to model the spectra of active regions. For this reason, we do not recommend that users adopt the IAG spectra as input of SOAP-GPU as long as we do

not have good observations of spots and faculae to use for modeling active regions. It is therefore crucial in the future to obtain active region observations at different heliocentric positions covering the full visible spectral range. Future solar observations using the ESPRESSO spectrograph may help us in that direction (Santos 2023).

After we filtered out lines with negative Pearson correlation coefficient values, we found that the line similarity of the Doppler component decreases with wavelength, which is likely due to the contamination of the telluric lines in the HARPS-N solar spectra. The line similarity of the Doppler component also increases with the line relative depth. This strong correlation is likely due to the low S/N of weak lines in the HARPS-N solar spectra that decreases the observed correlation (e.g., Cretignier et al. 2020). In the case of the line depth component, the line similarity also shows some correlation with relative line depth and the same explanation related to low S/N can be given as well.

We are conscious that the use of a single spectral line bisector to represent all of them is a current limitation to this first SOAP-GPU solar activity simulation at the spectral level, however, we demonstrated that even with this simplification, the general features of the HARPS-N solar spectra, both at the RV level, but also at the spectral line-shape variation level, can be well modeled by SOAP-GPU when using the SDO images as input. We are therefore confident that SOAP-GPU is able to satisfactorily model stellar activity on the Sun and therefore the published spectral time series should be considered as a useful test dataset for evaluating the performance of any stellar activity mitigating technique working at the spectral level. This does not guarantee that the same techniques working well on this dataset will work equally well on real observations as instrumental signals, but also other stellar signals such as supergranulation are not included in SOAP-GPU. This first SOAP-GPU simulation of solar activity at the spectral level could easily be improved further in terms of a realism once either IAG spectra for spots and faculae at different  $\mu$  angles, or MuRAM simulations modeling the full visible spectra of quiet and active regions at different  $\mu$  angles are available. It will then be easy to use those as input of SOAP-GPU to produce a more realistic solar simulation.

*Acknowledgements.* We thank the anonymous referee for the insightful and constructive feedback on this paper. X.D acknowledges the support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement SCORE No 851555) and from the Swiss National Science Foundation under the grant SPECTRE (No 200021\_215200). This work has been carried out within the framework of the NCCR PlanetS supported by the Swiss National Science Foundation under grants 51NF40\_182901 and 51NF40\_205606.

## References

Al Moulla, K., Dumusque, X., & Cretignier, M. 2024, A&A, 683, A106  
 Al Moulla, K., Dumusque, X., Cretignier, M., Zhao, Y., & Valenti, J. A. 2022, A&A, 664, A34  
 Al Moulla, K., Dumusque, X., Figueira, P., et al. 2023, A&A, 669, A39  
 Baumann, I. & Solanki, S. K. 2005, A&A, 443, 1061  
 Berdyugina, S. V. & Usoskin, I. G. 2003, A&A, 405, 1121  
 Borgniet, S., Meunier, N., & Lagrange, A. M. 2015, A&A, 581, A133  
 Bouchy, F., Pepe, F., & Queloz, D. 2001, A&A, 374, 733  
 Cavallini, F., Ceppatelli, G., & Righini, A. 1985, A&A, 143, 116  
 Collier Cameron, A., Ford, E. B., Shahaf, S., et al. 2021, MNRAS, 505, 1699  
 Collier Cameron, A., Mortier, A., Phillips, D., et al. 2019, MNRAS, 487, 1082  
 Crass, J., Gaudi, B. S., Leifer, S., et al. 2021, arXiv e-prints, arXiv:2107.14291  
 Cretignier, M., Dumusque, X., Allart, R., Pepe, F., & Lovis, C. 2020, A&A, 633, A76  
 Cretignier, M., Dumusque, X., Hara, N. C., & Pepe, F. 2021, A&A, 653, A43  
 Cretignier, M., Dumusque, X., & Pepe, F. 2022, A&A, 659, A68  
 de Beurs, Z. L., Vanderburg, A., Shallue, C. J., et al. 2022, AJ, 164, 49

Dumusque, X. 2018, A&A, 620, A47  
 Dumusque, X., Boisse, I., & Santos, N. C. 2014, ApJ, 796, 132  
 Dumusque, X., Cretignier, M., Sosnowska, D., et al. 2021, A&A, 648, A103  
 Dumusque, X., Glenday, A., Phillips, D. F., et al. 2015, ApJ, 814, L21  
 Ellwarth, M., Schäfer, S., Reiners, A., & Zechmeister, M. 2023, A&A, 673, A19  
 Feng, F., Tuomi, M., Jones, H. R. A., et al. 2017, AJ, 154, 135  
 Gilbertson, C., Ford, E. B., & Dumusque, X. 2020, Research Notes of the American Astronomical Society, 4, 59  
 Gray, D. F. 2009, ApJ, 697, 1032  
 Hathaway, D. H. 2011, Sol. Phys., 273, 221  
 Haywood, R. D., Collier Cameron, A., Unruh, Y. C., et al. 2016, MNRAS, 457, 3637  
 Husser, T. O., Wende-von Berg, S., Dreizler, S., et al. 2013, A&A, 553, A6  
 Lakeland, B. S., Naylor, T., Haywood, R. D., et al. 2024, MNRAS, 527, 7681  
 Langellier, N., Milbourne, T. W., Phillips, D. F., et al. 2021, AJ, 161, 287  
 Liebing, F., Jeffers, S. V., Reiners, A., & Zechmeister, M. 2021, A&A, 654, A168  
 Lin, A. S. J., Monson, A., Mahadevan, S., et al. 2022, AJ, 163, 184  
 Löhner-Böttcher, J., Schmidt, W., Schlichenmaier, R., Steinmetz, T., & Holzwarth, R. 2019, A&A, 624, A57  
 Martinez Pillet, V., Moreno-Insertis, F., & Vazquez, M. 1993, A&A, 274, 521  
 Mayor, M. & Queloz, D. 1995, Nature, 378, 355  
 Meunier, N., Desort, M., & Lagrange, A.-M. 2010, A&A, 512, A39  
 Palumbo, M. L., Ford, E. B., Gonzalez, E. B., et al. 2024, AJ, 168, 46  
 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  
 Rubenzahl, R. A., Halverson, S., Walawender, J., et al. 2023, PASP, 135, 125002  
 Santos, N. C. 2023, in Spectral Fidelity, 12  
 Schou, J., Scherrer, P. H., Bush, R. I., et al. 2012, Sol. Phys., 275, 229  
 SILSO World Data Center. 2024, International Sunspot Number Monthly Bulletin and online catalogue  
 Zhao, L. L., Dumusque, X., Ford, E. B., et al. 2023, AJ, 166, 173  
 Zhao, L. L., Fischer, D. A., Ford, E. B., et al. 2022, AJ, 163, 171  
 Zhao, Y. & Dumusque, X. 2023, A&A, 671, A11  
 Zhao, Y., Dumusque, X., Cretignier, M., et al. 2024, A&A, 687, A281

## Appendix A: Solar spectra modeling using IAG spectra

We summarize the results of solar activity simulations using IAG input seed spectra for the quiet solar disk. For modeling of active regions, we continue to use spectra from PHOENIX synthetic library with  $T_{\text{eff}} = 6028\text{K}$  for facula regions and  $T_{\text{eff}} = 5115\text{K}$  for spot regions. For both simulations, using either spot number or SDO data, we employ the same active regions map time series as in the PHOENIX case.

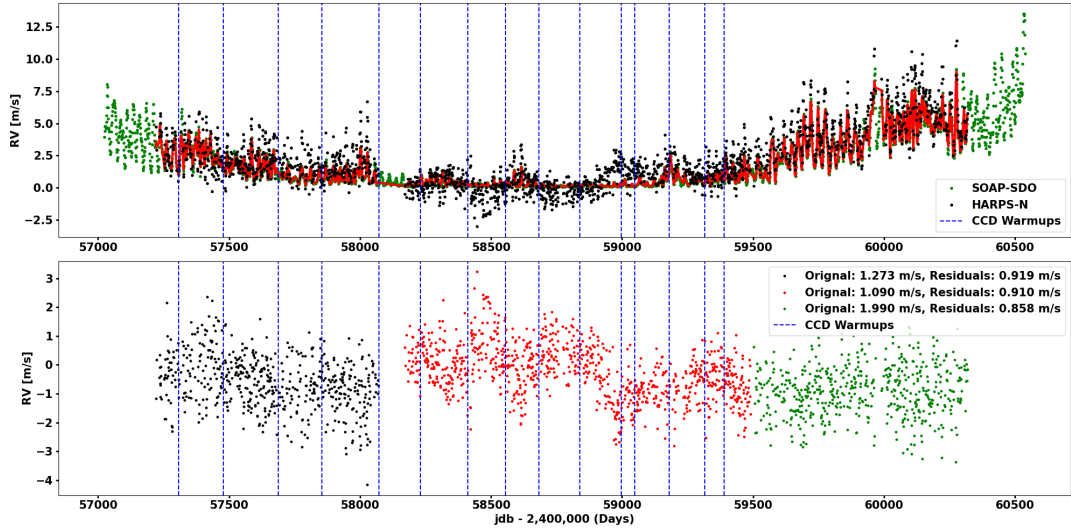


Fig. A.1. Same as Fig. 5, but simulated with the input seed spectra generated from the IAG observed solar spectra (see Sect. 2.2).

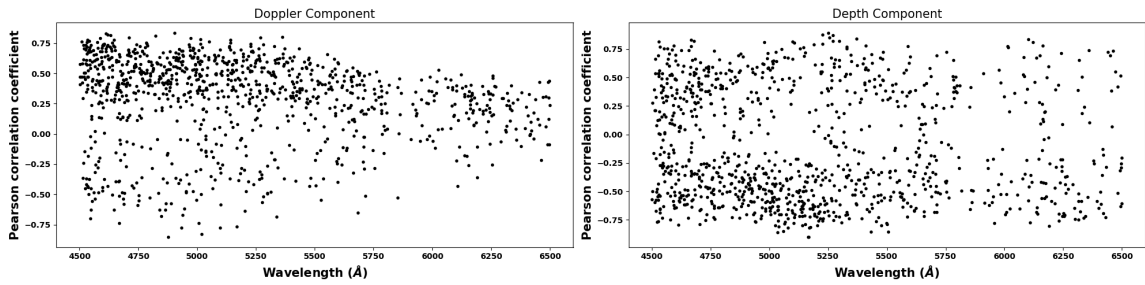


Fig. A.2. Same as Fig. 9, but simulated with the input seed spectra generated from the IAG observed solar spectra (see Sect. 2.2).

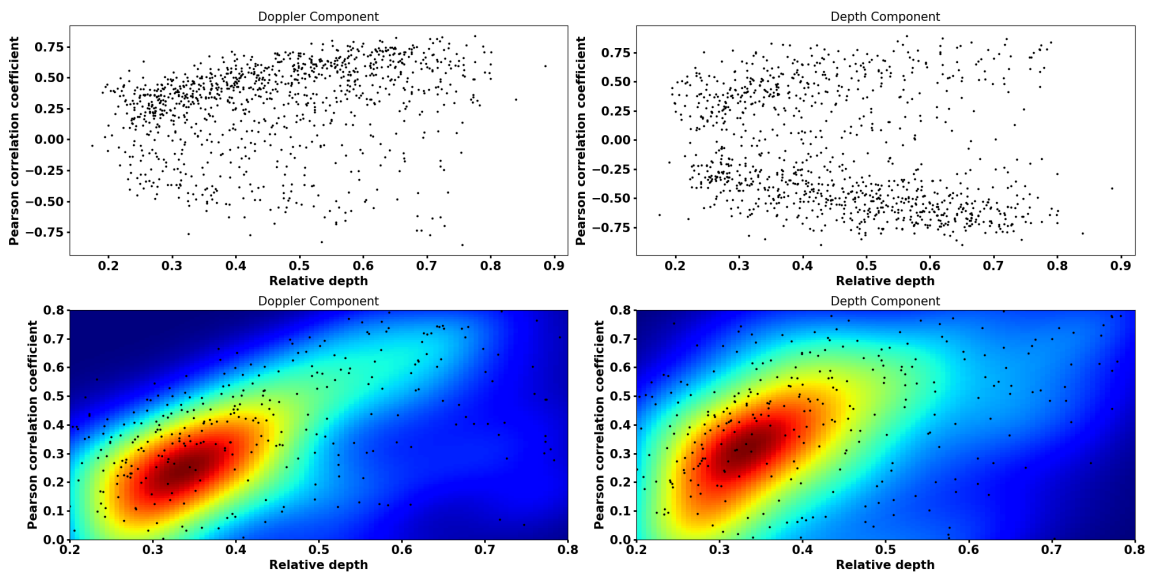


Fig. A.3. Same as Fig. 10, but simulated with the input seed spectra generated from the IAG observed solar spectra (see Sect. 2.2).