



Research

Cite this article: MacBride CD, Jess DB, Grant SDT, Khomenko E, Keys PH, Stangalini M. 2021 Accurately constraining velocity information from spectral imaging observations using machine learning techniques. *Phil. Trans. R. Soc. A* **379**: 20200171.
<https://doi.org/10.1098/rsta.2020.0171>

Accepted: 13 July 2020

One contribution of 16 to a Theo Murphy meeting issue 'High-resolution wave dynamics in the lower solar atmosphere'.

Subject Areas:

artificial intelligence, astrophysics, computational physics, observational astronomy, pattern recognition, solar system

Keywords:

methods: statistical, techniques: spectroscopic, Sun: atmosphere, Sun: chromosphere, Sun: photosphere, sunspots

Author for correspondence:

Conor D. MacBride
e-mail: cmacbride01@qub.ac.uk

Accurately constraining velocity information from spectral imaging observations using machine learning techniques

Conor D. MacBride¹, David B. Jess^{1,2},
Samuel D. T. Grant¹, Elena Khomenko^{3,4},
Peter H. Keys¹ and Marco Stangalini⁵

¹Astrophysics Research Centre, School of Mathematics and Physics, Queen's University Belfast, Belfast BT7 1NN, UK

²Department of Physics and Astronomy, California State University Northridge, Northridge, CA 91330, USA

³Instituto de Astrofísica de Canarias, 38205 La Laguna, Tenerife, Spain

⁴Departamento de Astrofísica, Universidad de La Laguna, 38205 La Laguna, Tenerife, Spain

⁵Italian Space Agency (ASI), Via del Politecnico snc, 00133 Roma, Italy

 CDM, 0000-0002-9901-8723; DBJ, 0000-0002-9155-8039;
SDTG, 0000-0001-5170-9747; EK, 0000-0003-3812-620X;
PHK, 0000-0001-8556-470X; MS, 0000-0002-5365-7546

Determining accurate plasma Doppler (line-of-sight) velocities from spectroscopic measurements is a challenging endeavour, especially when weak chromospheric absorption lines are often rapidly evolving and, hence, contain multiple spectral components in their constituent line profiles. Here, we present a novel method that employs machine learning techniques to identify the underlying components present within observed spectral lines, before subsequently constraining the constituent profiles through single or multiple Voigt fits. Our method allows active and quiescent components present in spectra to be identified and isolated for subsequent study. Lastly, we employ a Ca II 8542 Å spectral imaging dataset as a proof-of-concept study to benchmark the suitability of our code for

extracting two-component atmospheric profiles that are commonly present in sunspot chromospheres. Minimization tests are employed to validate the reliability of the results, achieving median reduced χ^2 -values equal to 1.03 between the observed and synthesized umbral line profiles.

This article is part of the Theo Murphy meeting issue 'High-resolution wave dynamics in the lower solar atmosphere'.

1. Introduction

Tomographic analysis of the complex solar atmosphere and its dynamics requires the combined use of different spectral lines (e.g. Ca II, Mg II, $\lambda\alpha$ and $\lambda\alpha$). Indeed, due to radiative transfer effects, these lines are formed across different heights of the solar atmosphere, with spectral information extracted at a particular percentage line depth corresponding to differing optical depths, and hence particular atmospheric heights [1]. Hence, it is possible to extract solar plasma properties as a function of optical depth (and atmospheric height if radiative transfer processes are well constrained) by taking measurements at specific percentage line depths [2].

The challenge of the solar chromosphere is that non-local thermodynamic equilibrium (non-LTE) plays an important role in the propagation of photons through the relatively dense and tenuous atmosphere [3]. Furthermore, due to spatial resolution limitations of even the largest-aperture ground-based solar telescopes, and the intrinsic fine-scale structuring of the solar atmosphere (scales smaller than 100–150 km), often the observed spectrum within the resolution element is the superposition of several components owing to different plasma states (e.g. both quiescent and magnetized atmospheric components). The effects of a multi-component atmosphere are generally taken into account in state-of-the-art spectropolarimetric inversion codes (e.g. NICOLE; [3]). However, these techniques are extremely demanding computationally. In addition to the effects of spatial resolution, multi-components can arise from the fact that chromospheric spectral lines sample a wide range of heights; from the mid-photosphere (wings of the line), through to the upper chromosphere (core of the line), thus mixing very different physical regimes [4].

Additional spectroscopic components that are superimposed on top of the quiescent background spectral profile can arise from a wealth of solar phenomena, including dynamic events such as magnetic reconnection, propagating waveforms, and shock development, all of which have the ability to form across a range of atmospheric heights. This adds a secondary component to an otherwise quiescent atmosphere [5], which is often highly Doppler-shifted relative to the approximately stationary line core, creating a significantly broadened resultant spectral profile. Being able to segregate the multi-component atmospheric contributions from a spectral line would allow the simultaneous examination of both 'dynamic' and 'quiescent' regimes of the atmosphere much more readily.

One of the most dynamic and widely recognized processes that creates multi-component chromospheric spectral lines is the propagation of magnetoacoustic waves, and their subsequent development into shock fronts, in the umbrae of sunspots [6–11]. Such umbral oscillations propagate energy upwards from the photosphere into the chromosphere, where the steep density drop results in the rapid increase of the velocity amplitude in an attempt to conserve energy flux, which is readily captured by the temperature-sensitive Ca II 8542 Å spectral line. Once the velocity amplitude exceeds the local sound speed, the waves develop into shock fronts—a process that can occur as low as approximately 250 km [12] through to the uppermost region of the chromosphere [11], resulting in the production of optically thin Ca II 8542 Å emission [13]. It is this optically thin emission that superimposes on top of the quiescent Ca II 8542 Å spectral profile, appearing in spectroscopic observations as an almost instantaneous blue-shifted (i.e. upwardly propagating) emission that slowly decays from its maximum intensity, providing the characteristic 'saw-tooth' spectral shape that is synonymous with umbral shock formation. Such spectral profile evolution

is coupled to the underlying *p*-mode wave spectrum, providing a periodicity of approximately 3 min to the umbral flashes [14,15]. Attempting to fit a multi-component umbral flash spectral profile using centre-of-gravity methods [16,17] or a single Gaussian, Lorentzian or Voigt profile is prone to error, since it may vastly under- or over-estimate the plasma characteristics of the developing shock. Hence, to fully understand the physics behind rapid atmospheric evolution (including those related to umbral flashes), both components must be considered separately; not simply modelled as a single-component atmosphere that attempts to bridge quiescent and active states [12,14].

Doppler (line-of-sight) velocities are typically acquired from spectroscopic observations using a number of techniques. They can be measured across a range of optical depths by calculating, at specific percentage line depths, the spectral line bisector shift from the central line core wavelength [2,18]. If full spectropolarimetric measurements are available, other atmospheric parameters (such as temperature, line-of-sight magnetic flux, and magnetic field inclination angles) can be estimated by inversion codes [19]. Such high-precision measurements of Stokes *I/Q/U/V* are difficult to obtain for deep chromospheric absorption lines, including Ca II 8542 Å, due to the relatively small rate at which photons reach the detector [20,21]. To combat the inherently weak signal, exposure times could be increased accordingly, but this is not always feasible when tracking the evolution of dynamic and rapidly evolving features in the lower solar atmosphere [22]. If only Stokes *I* spectral observations are available, Voigt profiles [23,24] can be fitted to the spectra. As we will discuss in §3, the Voigt profile includes enhanced spectral wing broadening, which provides a more suitable realization of the radiative transfer effects omnipresent throughout the lower solar atmosphere, including Doppler and pressure broadening mechanisms [25–28]. Alternative spectral shapes, such as the Gaussian and Lorentzian profiles, are useful when trying to represent profile shapes manifesting via different radiative transfer effects that are often present in the spectra.

Previous studies [12] have attempted to fit complex Ca II 8542 Å line profiles using a linear combination of two Gaussians, one for each component of the atmosphere attempting to be modelled. Boundary conditions for each fitted profile are employed such that the first Gaussian models an absorption component representative of the quiescent atmosphere, while the second Gaussian models an emission component linked to the dynamic phenomenon under investigation (e.g. magnetoacoustic shocks). This works well when two distinct components are present in an observed line profile, yet it introduces a number of challenges when only a single absorption component is resolved. Firstly, it takes more computational time and resources to fit the two-profile combination than a single profile, since twice the number of fitting parameters need to be constrained. Secondly, the boundary conditions placed on the fitted Gaussians tend to be subjectively chosen, and therefore may subject any extracted results to a user bias. Thirdly, the most problematic issue is the tendency for the two-profile combination to overfit the observed data as the fitting algorithm naturally attempts to smooth out noise and asymmetries in the observed spectra. This results in neither of the constituent profiles, by themselves, accurately modelling the absorption components present in the spectra, and Doppler shifts extracted from any one of the two fitted components would be potentially misleading and incorrect.

A novel method to alleviate the degree of human interaction required when processing large datasets hinges on the concept of machine learning, which is becoming increasingly commonplace in the field of solar physics. Machine learning, in particular neural networks, have been used in solar physics research from at least the early 1990s for applications such as predicting the maximum number of sunspots during a solar cycle [29], and also predicting the number of sunspots produced throughout the Sun's lifetime [30]. More recently, they have been used for predicting and detecting features on the Sun such as flares [31], coronal mass ejections [32] and active regions on the far side of the solar surface [33]. On small spatial scales, machine learning has been shown to play an important role in the investigation of velocity flows [34] and rapid spectropolarimetric inversions of Stokes *I/Q/U/V* spectra [35]. Hence, the application of machine learning and neural networks to the challenging problem of fitting complex spectral line profiles

(e.g. Ca II 8542 Å) is both timely and important for the accurate study of small-scale dynamic phenomena that are ubiquitously observed to permeate the solar atmosphere.

In this paper, we detail a method for fitting profiles to spectral lines that often contain multiple atmospheric components. Our method uses machine learning techniques to distinguish between spectra that have multi-component atmospheres and those that can be best represented by single spectral fits, allowing us to adjust the model and fitting method accordingly. We also provide a proof-of-concept study on a challenging Ca II 8542 Å spectral imaging dataset to show the suitability of this technique for widespread usage.

2. Observations

Spectral imaging observations of active region NOAA 12149 were acquired from 14:37–17:37 UT on 30 August 2014 using the Interferometric BIdimensional Spectrometer (IBIS; [36]) instrument at the National Science Foundation’s Dunn Solar Telescope, New Mexico, USA. IBIS is a Fabry–Pérot instrument able to obtain high temporal, spatial, and spectral resolution imaging spectroscopy measurements of the solar atmosphere. The Ca II 8542 Å spectral line was chosen for our observations, with IBIS operating in a Stokes *I* imaging mode (i.e. no polarimetric measurements were acquired) in order to maximize the temporal resolution of the spectroscopic scans. The observations obtained consisted of 2103 spectral scans using 27 non-equidistant wavelength points sampling over a 2.4 Å window centred on the Ca II 8542 Å line core. However, for the purposes of subsequent analysis, we focus on the central 23 wavelength points over a 1.6 Å window centred on the Ca II 8542 Å line core. As shown in figure 1, closer wavelength spacing was chosen around the line core to provide better quiescent profile fitting. The spatial sampling was 0''.098 per pixel, and the cadence of each Ca II 8542 Å scan was 5.8 s.

A contextual continuum image of active region NOAA 12149 was acquired from the Helioseismic and Magnetic Imager (HMI; [37]) onboard the Solar Dynamics Observatory (SDO; [38]). This measurement was taken at the start of the IBIS spectral imaging sequence, and once processed through standard SunPy data reduction algorithms [39,40] provided a full disk reference image with a spatial sampling equal to 0''.6 per pixel. Using the HMI contextual image to define absolute solar coordinates, the IBIS dataset was subsequently co-aligned to it using cross-correlation techniques [41]. Figure 2 shows example images from the IBIS dataset co-aligned with the SDO/HMI continuum.

3. Methods

Presented in this section are the details of a method which uses a neural network to classify a spectrum based on its profile shape. A suitable spectral fitting model is then selected for the spectrum according to its neural network classification. By accurate fitting of the chosen model, a Doppler shift, and therefore a velocity, can be uncovered for each of the constituent spectral components, thus allowing quiescent and dynamic parts of the atmosphere to be isolated and studied independently.

An average quiet Sun spectrum was extracted from the dataset by averaging all of the IBIS spectra contained within the rectangular region plotted in figure 2 across all 2103 scans. The average quiet Sun spectrum incorporated the averaging of 70 660 800 individual spectra, with the resulting profile plotted in figure 1. This location was chosen for constructing the quiet Sun spectrum because it is isolated from both the sunspot umbra (where complex Ca II 8542 Å profiles synonymous with shock formation are known to form) and the regions containing magnetic pores and bright points that often result in enhanced Ca II 8542 Å line core emission due to the formation of chromospheric plage [42–44]. The quiet Sun spectrum shown in figure 1 demonstrates a deep absorption profile with no clear evidence of enhanced wing or line-core emission, and hence forms a good ‘rest’ profile of the solar atmosphere that can be used to benchmark spectral fluctuations resulting from dynamic Ca II 8542 Å phenomena.

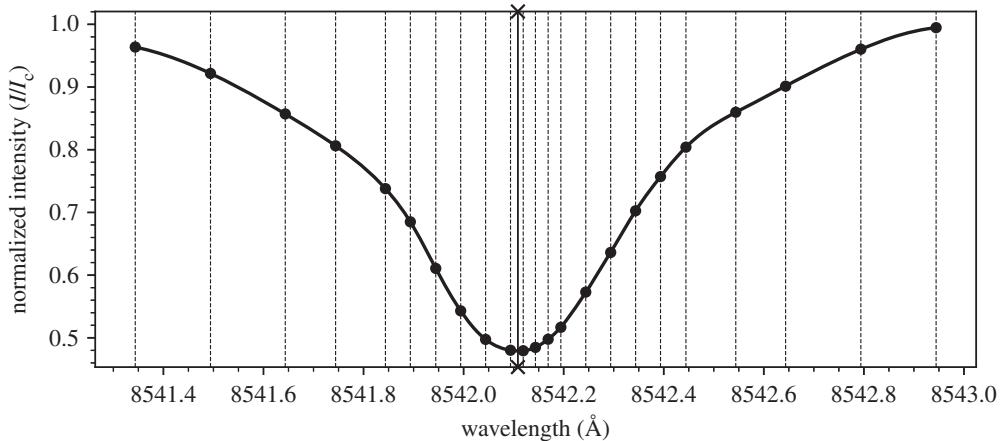


Figure 1. Plot of the average Ca ii 8542 Å spectrum taken from the quiet Sun region shown in figure 2, where the intensities, I , have been normalized by the quiet Sun continuum intensity, I_c . Solid black markers and vertical dashed lines highlight the 23 wavelength points used in our analysis out of the 27 wavelength points that were sampled by IBIS. The vertical solid line, emphasized by crosses, represents the stationary line core wavelength. Each spectral scan took 5.8 s to complete.

(a) Choosing a fitting profile

Observations of a spectral line corresponding to a specific atomic electron transition provide emission and absorption signatures over a range of wavelengths (i.e. not just an infinitely narrow feature at the wavelength associated with the transitioning electron). This increased width of the spectral line is due to a number of effects including Doppler broadening, pressure broadening, and Zeeman splitting [25–28]. These effects can be replicated closely by convolving a Gaussian function with a Lorentzian profile, producing a Voigt function, V , defined by [23]

$$V(x; A, \sigma, \gamma) = A \int_{-\infty}^{\infty} G(u; \sigma) L(x - u; \gamma) du, \quad (3.1)$$

where A is a constant scaling for the amplitude, $G(x; \sigma) = \exp(-x^2/(2\sigma^2))/(\sigma\sqrt{2\pi})$ is a Gaussian function centred at zero, and $L(x; \gamma) = \gamma/(\pi(x^2 + \gamma^2))$ is a Lorentzian function centred at zero. The parameters σ and γ are the standard deviations of the Gaussian and the half-width at half-maximum of the Lorentzian, respectively. The variable of integration is u . Since the Lorentzian and Gaussian functions are centred on zero, the variable x that is passed into the Voigt function must first be shifted by subtracting the wavelength of the line core, x_0 , such that the line core wavelength is at zero in the corresponding function. The x_0 parameter must be re-added to the fitted spectra in order to precisely account for Doppler-shifted plasma.

A linear combination of Voigt profiles was chosen to model the absorption and emission components of the spectra, hence producing a ‘double Voigt’ model. If the chosen spectrum shows only emission or absorption profile shapes (e.g. like shown in figure 1), then a ‘single Voigt’ model is selected, thus avoiding the computational and overfitment drawbacks associated with multiple component fits when not necessary. Note that a background intensity level is not included in the generated models computed from equation (3.1) as this will be calculated and subtracted from the spectra before fitting, as detailed in §3d(ii). By modelling the spectra with Voigt profiles instead of just simple Gaussian functions, we can fit the spectral lines more accurately as we are also including line wing broadening that mimics the observed spectral lines.

Single or double Voigt profiles are fitted iteratively to each spectrum until it is sufficiently similar to the observed line profile. During the fitting process, the employed model (‘single Voigt’ or ‘double Voigt’) is statistically evaluated numerous times to monitor the level of convergence between the synthetic and observed line profiles. It is therefore necessary for the selected model

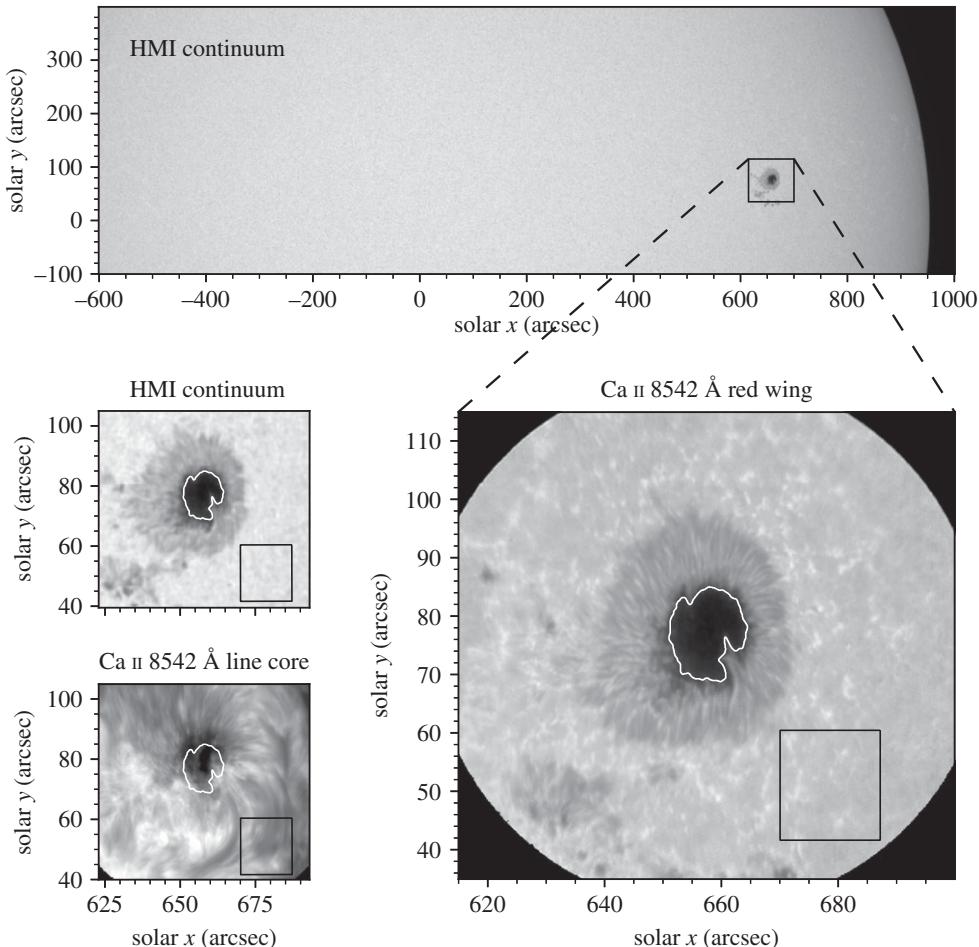


Figure 2. Co-aligned images of the IBIS Ca II 8542 Å red wing (line core $+1.2 \text{ \AA}$; lower right), IBIS Ca II 8542 Å line core (lower left), and SDO/HMI 6173 Å continuum (middle left) intensities. The location on the solar disk of active region NOAA 12149 during the observation period is shown in the top panel using a solid black box. The solid white contour used in each of the images represents the umbra-penumbra boundary used to isolate the umbral regions. The black rectangle displayed in each of the lower panels encloses a quiescent region of the solar atmosphere used to determine the average quiet Sun profile discussed in §3.

to be able to be evaluated efficiently, thereby saving computational time. The Voigt function was implemented in Python with the convolution being carried out by a numerical integration function, `quad`, provided by the `scipy.integrate` Python module [45,46]. An approximation of the Voigt profile [47] was also explored. However, even though this function requires less time to compute when compared to the numerical integration approach, the fitting methods exhibited very slow convergence, and hence we adopted the numerical integration technique for the remainder of our work.

We use a number of techniques to maximize the efficiency of our calculated Voigt functions. The integrand of the Voigt function is written in the C programming language and compiled into a shared library. This C function can then be imported into Python using the `ctypes` library. The `scipy.integrate` numerical integration function is then able to use this more efficient function as an integrand. We also adjust the absolute and relative error tolerances for the numerical integration, setting them to 0.149 and 0.000149, respectively. In practical terms, this means that the error in the raw intensity value at each wavelength point will be 0.149 for $|I - I_{BW}| \leq 1000$ and

$0.000149 \times |I - I_{\text{BW}}|$ for $|I - I_{\text{BW}}| > 1000$, where I are the intensity values at each wavelength point on a profile, and I_{BW} is the blue wing (rest line core – 1.2 Å) intensity value for the same profile.

Taking a typical full IBIS imaging spectral scan from our dataset, we computed the $|I - I_{\text{BW}}|$ value at every intensity value across all 664 796 individual spectra in the scan. We then calculated the error in each of the raw intensity values, I , using these computed $|I - I_{\text{BW}}|$ values. By dividing each error by I , a percentage error was determined for every intensity value. Over the entire field of view, the average percentage error was 0.00666%, with an error of 0.00739% over the subset of wavelength points ± 0.1 Å around the stationary line core. For the umbral spectra, which have much lower I_{BW} values, the average percentage error was 0.0179%, with an error of 0.0180% around the stationary line core. In our particular test, the maximum percentage error across the whole scan was only 0.0488%. Given that the IBIS spectra typically have line core and wing intensity values exceeding 1730 and 2550 detector counts, respectively, and $|I - I_{\text{BW}}|$ values of 757 and 149, respectively, these values correspond to absolute intensity errors of 0.149 detector counts. As such, these small percentage error tolerances are sufficient for our investigation.

(b) Neural network

(i) Classifying spectra

As discussed in §1, fitting a double Voigt model to all spectra introduces temporal efficiency and overfitting problems for those that contain only a single absorption/emission component. In these cases, a single Voigt model would be more appropriate. Our method for fitting spectral lines employs machine learning techniques to tailor the specific model (i.e. single Voigt or double Voigt functions) used for each spectrum sampled. This way, if a spectrum only contains a single component atmosphere consisting of an absorption dip or an emission peak, a single Voigt profile is fitted to the spectrum, with boundary constraints chosen objectively for such a line profile. On the other hand, if two components are detected, one absorption and one emission component, then a double Voigt profile is fitted, with each Voigt function appropriately constrained to fit the absorption and emission components of the observed line profile. In order to perform this automated task, we require a robust methodology to be able to reliably identify how many spectral components are present in each observed spectrum.

Machine learning, specifically artificial neural networks, are used to classify the spectra based on the amount of emission relative to absorption present in profile (see §3c for an in-depth discussion of how this is performed). Due to the diversity of spectral profiles in our dataset, it would have been very challenging to devise a fixed criterion that can be implemented into an algorithm and used to produce classifications more accurately than our neural network classifier. The neural network was trained to assign one of five classifications to each observed spectrum (figure 3), with the assigned classification used to adjust the model employed for final fitting. Specifically, the five classifications chosen consist of,

- (0) A profile exhibiting purely absorption characteristics, with no evidence of emission or asymmetric line wings.
- (1) An absorption line profile similar to the classification '0', but with evidence of a less precise line profile minimum due to, e.g. flattening of the line core intensities.
- (2) A profile that begins to show signs of embedded emission characteristics, yet the line core intensity is still less than the continuum intensity.
- (3) A line profile similar to classification '2', but with enhanced emission signatures such that the line core intensity is now just above the continuum level, and
- (4) A line profile showing heightened emission that is considerably brighter than the neighbouring continuum, hence dominating the resultant spectral profile.

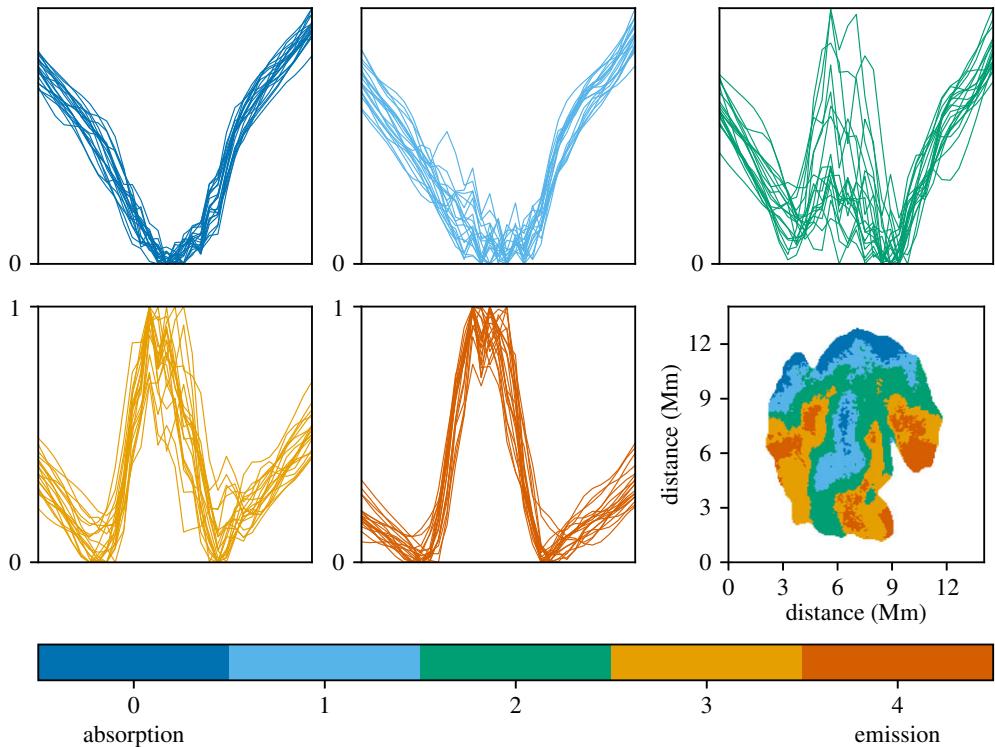


Figure 3. Plots of stacked Ca ii 8542 \AA umbral line spectra grouped by their neural network classification, where the intensity scale for each spectrum is normalized between 0 and 1 to aid visualization. A two-dimensional map (lower right) reveals the neural network classifications for the Ca ii 8542 \AA spectra present within the umbra (see contours in figure 2) for a single IBIS spectral imaging scan. (Online version in colour.)

This classification regime was chosen as it allows key features of the spectral shape to be identified and used to adapt the model, and it also identifies regions where the level of emission is particularly high. The neural network infrastructure is provided by the *scikit-learn* Python module [48–50]. From this package, we use the multi-layer perceptron (MLP) classifier [51] along with the L-BFGS-B solver [52], which is a unique version of the Broyden–Fletcher–Goldfarb–Shanno (BFGS) solver for limited memory and boundary constrained problems.

(ii) Mathematical processes

The observed spectrum is first interpolated on to a constant wavelength grid, resulting in a profile consisting of 33 data points. Our neural network implementation, with the multi-layer perceptron (MLP) classifier, takes the interpolated spectrum as a one-dimensional input layer with 33 neurons. The neural network then standardizes its input by rescaling the input vector to range from 0 to 1. As a general overview, each of the neurons of the input layer are subsequently connected to each of forty neurons on a single hidden layer. Finally, each neuron of the hidden layer is connected to all of the classifications on the five neuron output layer. The classification with the largest probability is assigned to the specific spectrum.

More specifically, each neuron–neuron connection between two layers has a weight, w_{ij} , where i is the index of the neuron in the current layer, and j is the index of the neuron on the previous layer. All of the neurons of the hidden and output layers have a bias, b . The input is propagated through the neural network along the connections between neurons of adjacent layers, with the

applied weights and biases determining the level of connectivity between the subsequent layers. The result at each neuron is calculated using the equation

$$y = g \left(\sum_j w_j x_j - b \right), \quad (3.2)$$

where x_j are the outputs from previous layers, w_j are their associated weights, and g represents the activation function.

The activation function for the neurons of the hidden layer was chosen to be the rectified linear unit function (ReLU; [53]) which is defined as $f(x) = \max(0, x)$, meaning that the neuron only ‘activates’ and passes a non-zero value on to the next layer if its result is positive. The output layer employed the softmax activation function, which is defined as $\text{softmax}(x)_i = \exp(x_i) / (\sum_{k=1}^5 \exp(x_k))$, where $i \in \{1, 2, 3, 4, 5\}$ are the five possible classifications that the neural network can assign.

This equation can be extended to apply to the output of a whole layer. Firstly, the outputs from the previous layer, x_j , can be represented as a column vector, $x \in \mathbb{R}^n$, where n is the number of neurons in the previous layer. Secondly, the weights, w_{ij} , can be represented as a matrix, $W \in \mathbb{R}^{m \times n}$, where m is the number of neurons in the current layer. Importantly, the matrix of weights between the input layer and the one hidden layer is given by $W^{(01)}$, and the input values are given by $x^{(0)}$. Similarly, $W^{(12)}$ represents the matrix of weights between the hidden layer and output layer, with $x^{(1)}$ representing the output of the hidden layer. The biases, $b \in \mathbb{R}^m$, for each neuron of the hidden layer are given by $b^{(1)}$, while the output layer can be represented by $b^{(2)}$. The output of the neural network is therefore $y \in \mathbb{R}^5$. As a result, a spectrum can be classified by evaluating the following set of equations in order:

$$x^{(1)} = f \left(W^{(01)} x^{(0)} + b^{(1)} \right) \quad (3.3)$$

and

$$y = \text{softmax} \left(W^{(12)} x^{(1)} + b^{(2)} \right), \quad (3.4)$$

where $f : \mathbb{R} \rightarrow \mathbb{R}$ and $\text{softmax} : \mathbb{R} \rightarrow \mathbb{R}$ are the ReLU and softmax activation functions, respectively, applied to the matrix elements. Subsequently, the spectrum is then assigned a classification associated with the largest output neuron value. Many spectra can be classified simultaneously by increasing the dimensions of the matrices forming tensors, which can be operated very efficiently using graphics processing units (GPUs; [54]).

The architecture of the neural network described in this study can classify a full IBIS imaging spectral scan, consisting of 664 796 individual spectra contained within the circular aperture of the instrument (figure 2), in approximately 48 seconds using a single 2.10 GHz Intel Xeon CPU core. As a result of this reasonable timeframe, porting to GPUs was not investigated in the present study. However, with larger field-of-view sizes and more rapid cadences from upcoming next-generation instrumentation [55], the resulting increase in data volume may require attention to be turned towards GPUs to provide the necessary accelerated performance to implement this technique in near real-time on next-generation datasets.

(c) Training procedure

In §3b, the weights and biases are initially unknown. For the neural network to be able to accurately classify spectra, the optimal weights and biases must be calculated, which is achieved by fitting the parameters of the neural network to a set of manually labelled training spectra.

The current accuracy of the neural network is quantified by a loss function, calculated using the manually assigned classification and the classification currently being assigned by the neural

network. The weights and biases are optimized by minimizing the value of the loss function. In this method, we use the log-loss function

$$\text{log-loss} = \frac{1}{N_S} \sum_{(y, y') \in S} (-y \log(y') - (1 - y) \log(1 - y')), \quad (3.5)$$

where S is the set of ground truth labels (y), defined as either '0' or '1', and currently predicted probabilities (y') spanning a range between 0 and 1, for each classification of the spectra in the training set. The parameter N_S is the number of (y, y') pairs in the set S .

To find the predicted probabilities for each classification, the algorithm passes the training set of spectra through the neural network, configured with an initially random set of weights and biases, before calculating the subsequent log-loss function. Finally, the L2 regularization term,

$$\text{L2} = \frac{\alpha}{2N_S} \left(\|W^{(01)}\|_2^2 + \|W^{(12)}\|_2^2 \right), \quad (3.6)$$

where $\|W\|_2^2$ is the sum of the weight matrices squared and α is the parameter that controls the amount of regularization, is added to the loss function. The training algorithm iteratively tries a number of different α terms, before selecting the specific value that produces the most accurate classifications after training.

The weights and biases that minimize the loss function are found using the L-BFGS-B quasi-Newton minimization algorithm [52], a variant of L-BFGS. This efficient algorithm works by finding an approximation of the Hessian matrix, which is a matrix of second derivatives of the loss function with respect to the weights and biases. Importantly, the Hessian matrix forms part of the quadratic term in the second-order Taylor expansion of the loss function when evaluated for a one-dimensional vector composed of the current weights and biases. The inverse of the Hessian matrix can also be found efficiently, which is then used to determine the step direction the vector of weights and biases should take for minimization of the loss function.

There are two separate sets of manually labelled ground truth data, one used for training the neural network and another used for testing its performance. By using a different set of spectra to test the performance of the neural network, we are able to detect if the neural network is overfitting, which would result in the test set being assigned less accurate classifications. Various statistics are determined, including a calculation of the percentage of spectra the neural network correctly assigned the ground truth classification.

The MLP classifier in the *scikit-learn* module also has a number of parameters that must be specified before the training is carried out. These include the numbers of hidden layers and neurons employed, the α parameter used for L2 regularization, the activation function used for the hidden layers (ReLU), and the algorithm used to optimize the weights and biases. A neural network consisting of a single hidden layer with 40 neurons was chosen for our study.

The L-BFGS, stochastic gradient descent (SGD) and Adam [56] optimization algorithms were trialled. Upon successive trainings, the L-BFGS method produced trained neural networks with the most consistent accuracy score, and hence this approach was adopted for the current study. The accuracy score is defined as the fraction of the testing set of spectra that the neural network successfully assigned its manually labelled classification. Our method uses the *GridSearchCV* infrastructure provided by the *scikit-learn* module to select the best α parameter from the list $[1, 2, 3, 4, 5, 6, 7, 8, 9] \times 10^{-5}$.

(i) Training with the IBIS dataset

To train and test the neural network, 200 spectra from the IBIS Ca II 8542 Å dataset introduced in §2 were chosen at random from within the umbra. The spectra in this set are manually labelled with a classification, providing ground truth data for training and testing purposes. The first half of this 200 spectra dataset is used to train the neural network, while the other half is used to test the success of the trained neural network. A subset of these spectra are separated into their manually labelled classifications and plotted in figure 3 for visual clarity. As described in §3b(i),

these spectra are classified based on the ratio of emission to absorption, with the extreme ends of the scale being those spectra showing no evidence of emission (classified as '0'), and those spectra with obvious emission much brighter than the neighbouring continuum (classified as '4').

Following the methodology outlined in §3c, the success of the trained neural network can be classified in terms of 'precision' and 'recall' statistics. Here, the precision is defined as the percentage of spectra that were correctly assigned a classification, c , from all the spectra that were also assigned the same classification by the neural network. Recall is defined as the percentage of spectra that were correctly assigned a classification, c , out of all the spectra that should have been assigned this specific classification according the ground truth data. For each of the classifications ($0 \rightarrow 4$) in ascending order, the precisions obtained were 77%, 74%, 71%, 100% and 94%. Similarly, in terms of recall, the percentages obtained were 94%, 78%, 80%, 62% and 100%. These result in overall averages, weighted by the number of ground truth spectra per classification, of 83% and 81% for the precision and recall parameters, respectively.

Importantly, 95% of the spectra incorrectly classified by the neural network had a deviation of only one classification number. These performance measures suggest that the neural network is sufficiently accurate for our purposes. Running the neural network classification algorithm on the umbral pixels extracted from a single IBIS spectral imaging scan displays the expected complexities associated with this type of solar feature. In particular, the lower right panel of figure 3 displays the classifications that the trained neural network returned, revealing a multitude of purely absorption (i.e. quiescent) and emission (i.e. active) spectra.

(d) Fitting method

The following preprocessing steps and fitting methods are applied to each spectrum of the IBIS dataset independently. Based on the neural network processing described in §3c(i), each spectrum is assigned a classification. Spectra that are assigned a classification of either '0' or '1' are fitted with a simple, single Voigt profile as described in §3a. Furthermore, spectra assigned a classification of either '2', '3' or '4' are modelled with double Voigt profiles. This results in spectra with a dominant absorption component being fitted with a single quiescent atmosphere model, whereas spectra with increasing emission components are modelled with multi-component fits that are more representative of the dynamic atmosphere. Doppler velocity information can then be extracted for both the quiescent and active atmospheres using the parameters inferred from the fitted models.

As will be discussed in §3d(iii), we will further distinguish between classifications '0' and $1 \rightarrow 4$. However, classifications $2 \rightarrow 4$ are processed in an identical manner. While we found no additional benefits to distinguishing between these different classifications when fitting, we have kept these as distinct classifications to demonstrate the adaptability of our method and to assist with relating the spatial distribution of classifications to any dynamic features found in the Doppler velocity measurements.

(i) Preprocessing techniques

As shown in figure 1, the spectral images acquired using IBIS consist of intensity measurements obtained across 27 non-equidistant wavelength points. However, in our analysis we focus on the central 23 wavelength points, providing a 1.6 \AA wavelength range across the Ca II line, allowing Doppler velocities spanning $\approx \pm 30 \text{ km/s}$ to be studied. Each spectrum is interpolated on to an equidistant wavelength grid employing a wavelength spacing of 0.05 \AA . This produced a spectrum comprising 33 equidistant wavelength samples.

The next stage is to ensure the line wings are correctly calibrated. IBIS includes a prefilter with an FWHM of 4.6 \AA centred at 8542.5 \AA in order to isolate the 8542 \AA signal from the other resonant wavelengths produced by the Fabry-Pérot interferometer [26]. As a result, each profile is initially corrected by removing the prefilter transmission response from the measured spectrum.

(ii) Background removal

In the fitted models, a background profile is not included, and as a result the continuum intensity is expected to be zero. The motivation for not including a background profile is to reduce the number of free parameters, and thus make the model less computationally challenging to fit. Hence, a constant background intensity is calculated for each spectrum by performing a boxcar moving average of the intensities calculated 1.2 Å into the blue wing from the Ca II 8542 Å line core over a time frame of ± 2 min relative to the current scan, which is subtracted from the spectrum before fitting with a suitable model. When processing a number of spectra over a large field of view, the calculation of the backgrounds and their subsequent subtraction can be vectorized, resulting in a significant computational speed improvement compared to processing each spectrum individually. Due to the breadth of the Ca II 8542 Å line, in conjunction with the relatively narrow wavelength range of the IBIS prefilter, it is not possible to calculate the true continuum intensity for our dataset. As such, we approximate the continuum intensity, I_c , using a boxcar moving average placed at the furthest blue-wing position of the IBIS spectral imaging scans (i.e. line core -1.2 Å).

(iii) Weighted fit

In order to fit the chosen model to the spectra more accurately, we applied weights to different parts of the spectrum. Weights are given by assigning error bars to the spectrum, which the fitting algorithm uses when calculating goodness-of-fit parameters. As the measurement of Doppler velocities is of paramount importance, especially for quiescent spectra classified as '0' where the Doppler shifts may be subtle, we prioritize fitting around the spectral line core. As a result, we add larger uncertainties to the wings of the spectral line when compared to the line core, which can be visualized in the upper panel of figure 4.

On the other hand, for spectral classifications 1 \rightarrow 4, we found that the accuracy of the fit was improved by increasing the associated uncertainties at wavelengths immediately surrounding the wavelength of the stationary line core (see the lower panel of figure 4). This had the effect of lowering the priority for fitting the exact shape of the peak, which was either difficult to isolate (e.g. for classifications '1' and '2' shown in figure 3), or incentivized the optimization algorithms to fit more closely the steepening gradients of the spectral line wings (e.g. for classifications '3' and '4' shown in figure 3) that are representative of the dynamic activity present in the spectra.

(iv) Least-squares optimization

The chosen model (i.e. either a single or double Voigt model) is fitted to the spectra using the `curve_fit` function provided by the `scipy.optimize` Python module [45,46]. With this function we use the Trust Region Reflective algorithm for least-squares optimization, similar to that proposed by Branch *et al.* [57], which allows for bounds to be specified for the parameters intrinsic to the model, which assists with spectral convergence.

The amplitude is bounded such that a single Voigt model must fit a negative (i.e. absorption) amplitude, while a double Voigt model must fit one negative amplitude (representing the quiescent atmosphere) and one positive amplitude (representing the active atmospheric component). This helps to prevent overfitting issues described in §1.

The wavelength at the centre of the absorption Voigt profile is constrained such that it does not deviate too far from the stationary line core wavelength. The central wavelength of the absorption Voigt function is allowed to vary within a 0.15 Å window above and below the stationary line core wavelength. The stationary line core wavelength is found by applying this method, without bounds, to the average quiet Sun spectrum introduced in §2. This constrains the method to only allow Doppler velocities in the absorption component that are ± 5.27 km s $^{-1}$. This helps ensure that the model represents what is happening physically in the spectra, and does not give an unphysical Doppler shift.

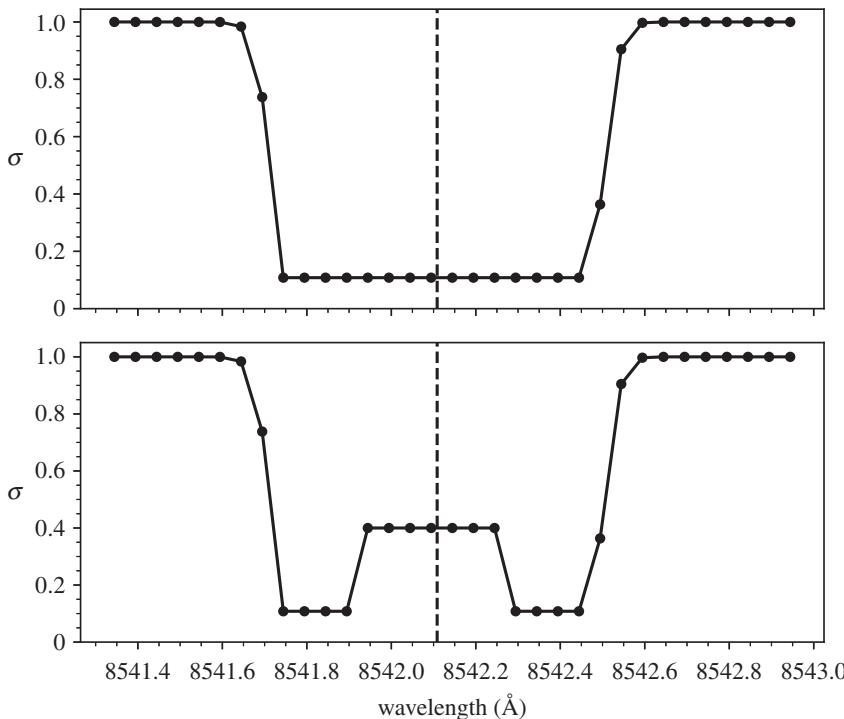


Figure 4. Plots of the sigma profiles used to weight particular regions of the spectra during the fitting process. The top profile is used for spectra of classification '0' and has the effect of reducing the priority of fitting the wings. The bottom profile is used for all other classifications of spectra and reduces the priority of fitting the precise shape of the spectral turning point. The vertical dashed line represents the stationary line core wavelength that these profiles are mapped on to.

Bounds are also placed on the γ and σ parameters for the Lorentzian and Gaussian shapes that are convolved to form the resultant Voigt profile. The bounds chosen for the γ and σ components are $10^{-6} < \gamma$ or $\sigma < 1$, which allows a wide range of spectral broadening to be accounted for. An initial guess is also supplied to the fitting function, where the central wavelength of each constituent Voigt profile is initially set to the stationary line core wavelength. Any Voigt profiles modelling absorption are given a negative initial amplitude that is typical for the spectra dataset, while any Voigt profiles representing emission are given a typical positive initial amplitude. All γ and σ initial guesses are taken to be 0.1 and 0.2, respectively.

Figure 5 shows examples of the fitting methods applied to IBIS spectra. In the left panel, the method is applied to a spectrum containing only an absorption component and is therefore fitted with a single Voigt model. A spectrum exhibiting a complex mix of absorption and emission is fitted with a double Voigt model in the right panel. Both of these fitted profiles show excellent agreement with the observed profiles, in particular, around the line core where the fitting was prioritized.

(e) Calculating velocities

Once the quiescent and active components have been isolated, they can be studied independently to ascertain their respective Doppler shifts. The wavelength of the line core, $\lambda_{\text{observed}}$, for each atmospheric component can be extracted from the parameters defining the central wavelengths of the best-fitting constituent Voigt profiles. The 'at rest' line core wavelength, $\lambda_{\text{stationary}}$, is calculated by applying the algorithm to a spatially and temporally averaged spectrum extracted from a region containing quiet Sun, as documented in §2 (see, e.g. the rectangular quiet Sun region

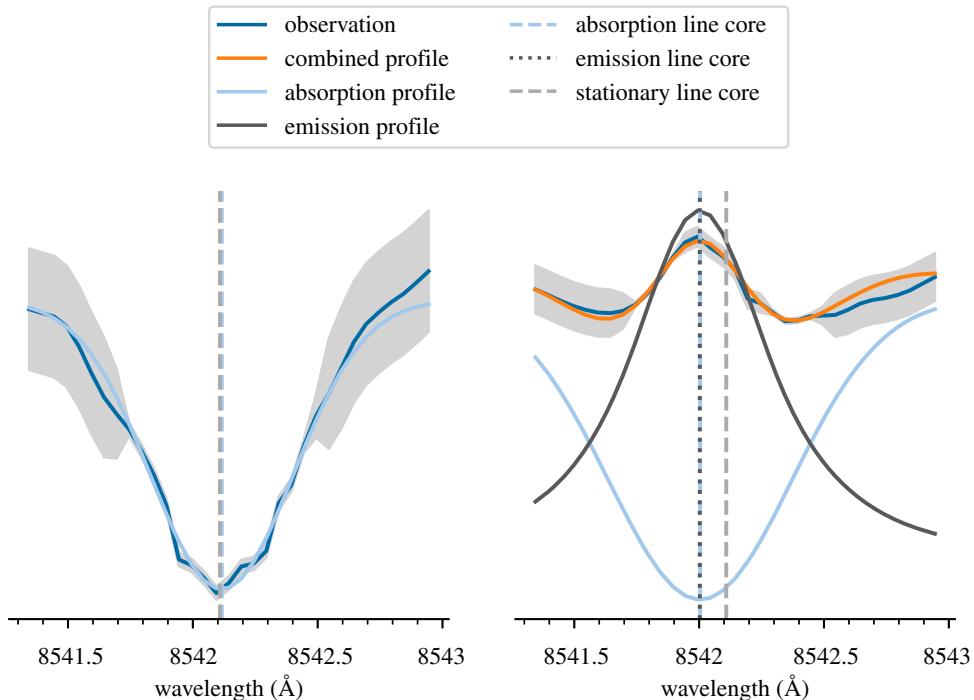


Figure 5. Examples showing the fitting methods applied to spectra with a single component atmosphere (neural network classification '0'; left panel) and a multi-component atmosphere (neural network classification '3'; right panel). The labelled vertical lines identify the line core wavelengths of each atmospheric component fitted, i.e. the central wavelengths of the relevant Voigt functions. The addition of the multi-component fits (right panel) shows close agreement with the observed profile. The grey shaded regions plotted on top of the spectra represent the sigma weighting profiles displayed in figure 4. (Online version in colour.)

depicted in figure 2 and its average profile shown in figure 1). Doppler velocities, v , can then be calculated by comparing the line core wavelengths of each of the atmospheric components to the stationary wavelength via,

$$v \text{ (km s}^{-1}\text{)} = \frac{\lambda_{\text{observed}} - \lambda_{\text{stationary}}}{\lambda_{\text{stationary}}} \times 300\,000 \quad (3.7)$$

4. Proof of concept testing with IBIS data

A proof of concept test was performed using the IBIS dataset described in §2. The spectra across the full field of view for a single spectral imaging scan (totalling 664 796 individual spectra) were first classified by the neural network, with the resulting classifications shown in figure 6. Following the methods outlined in §3, Doppler velocities were computed for each fitted absorption component, with the resulting velocity maps shown in the left panel of figure 7. If a spectral profile required multiple profile fits (i.e. using a double Voigt model), then the resulting Doppler velocities inferred from the emission Voigt fit are shown in the right panel of figure 7. In particular, it can be seen in the right panel of figure 7 that many of the derived Doppler velocities associated with dynamic phenomena appear to rapidly change between neighbouring pixels, an effect that has been documented in previous velocity studies of the solar chromosphere [58–60]. These rapid velocity excursions appear to be closely linked to instances when the Ca II 8542 Å line goes into emission. As a result, we believe that such discontinuities may be caused by dynamic changes in the opacity of the plasma, resulting in shifts of the response function of the line

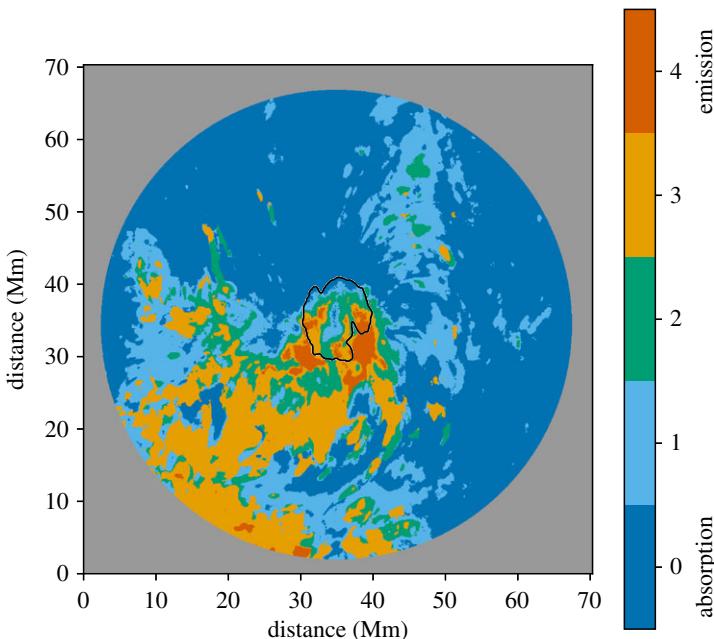


Figure 6. The neural network classifications of the spectra extracted from a single IBIS spectral imaging scan, where the colour bar relates to the spectral shape classified, with '0' and '4' representing pure absorption and emission profiles, respectively. The umbra/penumbra boundary is highlighted using a black contour, and is consistent with that shown in figure 2. (Online version in colour.)

caused by the source function no longer monotonically decreasing throughout the chromosphere [9,61], and hence are not numeric artefacts. As such, we must stress that smooth velocity fields should not necessarily be expected following the application of our techniques, especially when examining the challenging effects of radiative transfer in the solar chromosphere.

A goodness-of-fit value was estimated at each pixel of the field of view using a modified χ^2 relation

$$\chi^2 = \frac{s}{\nu} \sum_{\lambda \in \lambda_c} \frac{(I_\lambda^{\text{fitted}} - I_\lambda^{\text{observed}})^2}{I_\lambda^{\text{observed}}}, \quad (4.1)$$

where s is a scaling factor, ν is the estimated degrees of freedom, and $I_\lambda^{\text{fitted}}$ and $I_\lambda^{\text{observed}}$ are the intensity values of the fitted and observed spectra, respectively. The wavelengths over which this calculation is performed, λ_c , include the wavelength closest to the stationary line core, in addition to the 12 wavelength points either side. Thus, the central 25 wavelength points (out of 33 points in total) were used to compute the modified χ^2 -statistic as this allowed for a better measure of the goodness-of-fit around the line core. Therefore, we introduce the scaling factor, $s = \frac{33}{25}$, to account for this subset of wavelength points. The degrees of freedom, ν , were calculated for each spectrum by subtracting the number of fitted parameters (4 for single Voigt profiles and 8 for double Voigt profiles) from the total number of wavelength points, 33. The χ^2 -values for all the fitted spectra in the field of view are summarized in figure 8.

Considering the entire IBIS spectral imaging scan, the median χ^2 -value is $\tilde{\chi}^2 = 2.36$. When only the modified χ^2 -values for the umbral locations (i.e. spectra within the white umbral contours highlighted in figure 2) are used, the median value is $\tilde{\chi}^2 = 1.03$. The calculated median χ^2 -values are close to one, suggesting that the fitting method is able to accurately constrain the observed spectral line profiles. The accuracy is particularly good when considering umbral pixels, since

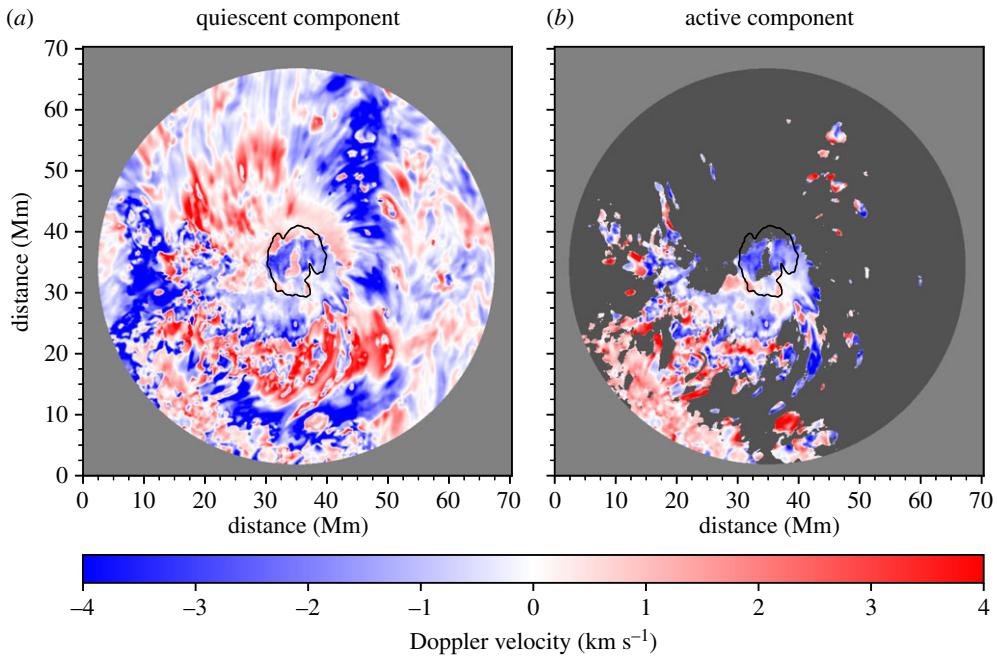


Figure 7. The Doppler velocities corresponding to the fitted absorption components are shown in the left panel, with the colour bar saturated between $\pm 4 \text{ km s}^{-1}$ to aid visual clarity. For those profiles identified as containing two atmospheric components (i.e. are fitted using a double Voigt model), the Doppler velocities of the secondary emission profile are shown in the right panel. Pixels not requiring a secondary profile fit are shown with a darker grey colour to better isolate them from those pixels requiring a double Voigt model. (Online version in colour.)

not only is the median χ^2 -value particularly close to one, as can be seen in figure 8, but the tail on the χ^2 -distribution for umbral spectra drops off very rapidly with increasing χ^2 -values.

The most computationally intensive aspect of the proof of concept testing was the fitting of suitable Voigt models to the spectra, with the double Voigt model taking more time than the single Voigt model. To fit a full spectral imaging scan totalling 664 796 individual spectra, with 30% being modelled using a double Voigt profile (figure 6), took 123 min running across all 16 cores on a 2.10 GHz Intel Xeon processor. As a result, processing all 2103 spectral scans from the current IBIS dataset on a single CPU would likely take on the order of 175 days. However, the techniques presented can be further parallelized by employing multiple CPUs, hence bringing the entire processing time down to the order of one week or less. Furthermore, as discussed in §3b(ii), the current algorithms may be ported across to GPUs, providing the ability to accelerate processing performance by an order of magnitude or more.

5. Discussion

Although our primary objective for this method is to accurately constrain velocity information within an umbral region, the method can be applied more generally to any region of a spectral imaging dataset where a two-component atmosphere may be present. As shown in our proof of concept test (§4), our methods can be applied to any solar region, including both dynamic locations (umbrae, penumbrae, regions of magnetism, etc.) and those demonstrating more quiescent behaviour (e.g. quiet Sun that is permeated by granulation). Although the set of labelled ground truth spectra that are used to train and test the neural network are chosen from within the umbral region, the range of spectral profile shapes encountered within the umbra (i.e. spanning

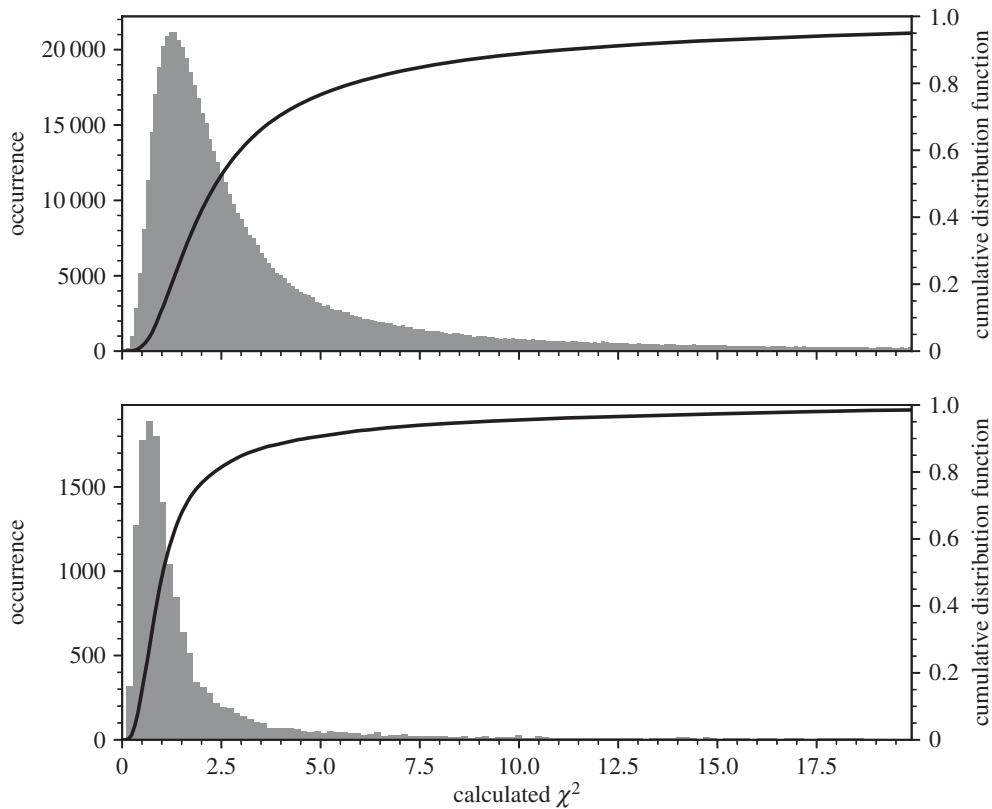


Figure 8. A histogram of occurrences of the modified χ^2 -values (derived using equation (4.1)) for a single IBIS spectral imaging scan (top panel) is plotted in grey, which depicts the goodness-of-fits between the modelled and observed spectral profiles. In the bottom panel, the histogram is reproduced including only χ^2 -values for umbral pixels (as highlighted by the white contours in figure 2). In each panel, the solid black lines show the cumulative distribution functions, with the highest 5.0% and 1.5% of the calculated χ^2 -values omitted from the top and bottom panels, respectively. The median χ^2 -values for the entire field of view and the umbral locations are $\tilde{\chi}^2 = 2.36$ and $\tilde{\chi}^2 = 1.03$, respectively.

pure absorption through to pure emission characteristics) are representative of many different spectra found outside of the umbral region.

The precision and recall scores of 83% and 81%, respectively, as introduced in §3c(i), suggest that the neural network is able to classify spectra with a reasonable degree of accuracy. Since our method currently treats classifications '2', '3' and '4' identically, the performance statistics can be recalculated assuming these cases all have the same classification. This results in increased precision and recall scores of 91% and 90%, respectively, suggesting that the performance of the neural network is particularly well suited for our methods. Similarly, if we adjust the neural network to distinguish between 'emission' (classifications '2', '3' and '4') and 'no emission' (classifications '0' and '1'), the precision and recall scores for these groupings are 96% and 95%, respectively. With a larger set of ground truth data, perhaps including other highly dynamical solar phenomena that often exhibit enhanced line-wing asymmetries (e.g. penumbral jets, spicules, magnetic reconnection; [62–64]), the precision and recall scores of the neural network could be improved yet further, even to be very close to 100%. The precision and recall values computed here are consistent with other trained astrophysical neural networks adopted into mainstream data processing [65,66].

The classification methods presented here (i.e. excluding the line fitment and velocity processing) may also be useful for estimating the degree of quiescence in a particular dataset,

where the pixel fractions related to ‘quiescent’ and ‘active’ components can be readily compared. A similar methodology has already been applied to optimize the inversion of the He I 10830 Å line [67], and the classification aspects of this code can be implemented into future inversions of Ca II 8542 Å data with tools such as the Non-LTE Inversion Code using the Lorien Engine (NICOLE; [3]). Such processing is carried out by applying only the neural network classification procedure to the data and monitoring the relative occurrence of each of the five classifications—a process that can be accomplished within a few minutes.

In the future, our methods could be adapted to find velocity measurements for chromospheric jets that often demonstrate plasma motions in the range of approximately 20–40 km s^{−1} [68]. Another potential use case is the study of Ellerman bombs, which have been observed in the Ca II 8542 Å line core as significant blue-shifted emission [69,70]. This blue-shifted emission is present in other chromospheric lines, including H α and He I 10830 Å, but not in any photospheric lines [70]. As described above, it may be necessary to further train the neural network using a larger set of ground truth data, including examples of enhanced line-wing asymmetries that are synonymous with such dynamical solar phenomena.

Our method could also be extended to model a three (or more) component atmosphere by including additional Voigt (or similar) profiles in the model and modifying the criteria that determine how the assigned classification adjusts the fitting method. For such a technique to produce accurate velocities, a much higher number of wavelength samples would be required, such that the components are more clearly resolved and are therefore less blended with the surrounding plasma. With a higher number of wavelength samples, other features, such as the presence of double-peaked self-reversal structures in the emission components, could also be resolved. To facilitate such future code development, attention will likely need to be turned to the next generation of spectral imaging Fabry-Pérot instruments, in addition to slit- and fibre-based spectropolarimeters. These revolutionary instruments, including the Visible Tunable Filter (VTF; [71]) and the Diffraction Limited Near Infrared Spectropolarimeter (DL-NIRSP), will soon be commissioned by the National Science Foundation’s Daniel K. Inouye Solar Telescope (DKIST; [72]).

The ‘active’ and ‘quiescent’ components present in Stokes *I* observations can be isolated using our method (e.g. figure 7). Future work to investigate passing these isolated components through the Non-LTE Inversion Code using the Lorien Engine (NICOLE; [3]) or the CAlcium Inversion using a Spectral ARchive (CAISAR; [73]) inversion codes would be of particular interest. Often these inversion codes provide plasma parameter outputs based on a static atmosphere. Hence, by treating the two atmospheric components separately, ambiguities in the inverted plasma parameters could be minimized, since each component would be better constrained by its own single, high-quality spectral fit. This is a timely endeavour, especially with new, high spectral precision observations on the horizon from new telescope facilities, including DKIST.

The code is fully available for the community to download and use. Details of how to access it are provided in appendix A.

6. Conclusion

Novel methods for accurately constraining velocity information from spectral imaging observations have been presented. Using machine learning techniques, our methods automatically adapt the spectral models used to fit the input spectra. Importantly, our methods will only fit multi-component models if multiple signatures are observed in the input spectra, hence saving time and preventing the overfitment of the data.

By modelling each atmospheric component with its own independent Voigt function, the constituent components of the atmosphere can be isolated, both spatially and temporally. Such techniques have a diverse range of use cases, including the applicability to upcoming DKIST observations, as well as refining the inputs for modern inversion routines, since it enables each atmospheric component to be studied independently.

A proof of concept test applying this method to a challenging Ca II 8542 Å spectral imaging dataset was presented. In this, we demonstrated both the accuracy of the method and how its techniques can be applied more generally. Importantly, the algorithms presented are available to the global community through regularly updated download links.

Data accessibility. The data used in this paper are from the observing campaign entitled 'Nanoflare Activity in the Lower Solar Atmosphere' (NSO-SP proposal T1020; principal investigator: DBJ), which employed the ground-based Dunn Solar Telescope, USA, during August 2014. The Dunn Solar Telescope at Sacramento Peak/NM was operated by the National Solar Observatory (NSO). NSO is operated by the Association of Universities for Research in Astronomy (AURA), Inc., under cooperative agreement with the National Science Foundation (NSF). Additional supporting observations were obtained from the publicly available NASA's Solar Dynamics Observatory (<https://sdo.gsfc.nasa.gov>) data archive, which can be accessed via <http://jsoc.stanford.edu/ajax/lookdata.html>. The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

Authors' contributions. C.D.M. and D.B.J. conceived of and designed the study. D.B.J. carried out the experiments. C.D.M. performed the data reduction and scientific analysis, with assistance from D.B.J., S.D.T.G., P.H.K. and M.S. C.D.M. drafted the manuscript, with theoretical input provided by E.K. All authors read and approved the manuscript.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by The Department for the Economy (Northern Ireland) through their postgraduate research studentship and an Invest NI and Randox Laboratories Ltd. Research & Development (grant no. 059RDEN-1).

Acknowledgements. C.D.M. would like to thank the Northern Ireland Department for the Economy for the award of a PhD studentship. D.B.J. and S.D.T.G. are grateful to Invest NI and Randox Laboratories Ltd. for the award of a Research & Development that allowed the computational techniques employed to be developed. The authors wish to acknowledge scientific discussions with the Waves in the Lower Solar Atmosphere (WaLSA; www.WaLSA.team) team, which is supported by the Research Council of Norway (project no. 262622), and The Royal Society through the award of funding to host the Theo Murphy Discussion Meeting 'High-resolution wave dynamics in the lower solar atmosphere' (grant no. Hooke18b/SCTM).

Appendix A. Availability of the code

The methods presented in this study have been compiled into a Python package *Multi-Component Atmospheric Line Fitting* (MCALF; [74]).¹ This software package includes a number of subpackages, namely,

- `mcalf.models`: a collection of models, built upon a common base model, for fitting specific datasets;
- `mcalf.profiles`: profile functions that are used to form models;
- `mcalf.visualization`: a collection of functions for visualizing input and output data; and
- `mcalf.utils`: a collection of utility functions used throughout.

This package provides a 'toolkit' that can be used to define a model optimized for a particular dataset. A base model is provided that is suitable for any spectral imaging instrument, as well as the custom model derived in the present study that has been optimized for the IBIS sunspot dataset we introduce in §2. The user can easily build upon the base model using our custom template as an example. This allows the user to apply specific bounds on the fitted parameters and provide their own ground truth dataset of spectral shapes that they would like their neural network to be able to distinguish between. Additional logic can also be included to interpret their own neural network classifications and take care of the specific spectral shapes present in their dataset.

¹<https://github.com/ConorMacBride/mcalf>.

References

1. Vernazza JE, Avrett EH, Loeser R. 1981 Structure of the solar chromosphere. III - Models of the EUV brightness components of the quiet-sun. *Astrophys. J. Suppl. Ser.* **45**, 635. (doi:10.1086/190731)
2. González Manrique SJ, Quintero Noda C, Kuckein C, Ruiz Cobo B, Carlsson M. 2020 Capabilities of bisector analysis of the Si I 10827 Å line for estimating line-of-sight velocities in the quiet Sun. *Astron. Astrophys.* **634**, A19. (doi:10.1051/0004-6361/201937274)
3. Socas-Navarro H, Asensio Ramos A, Trujillo Bueno J, Ruiz Cobo B. 2015 An open-source, massively parallel code for non-LTE synthesis and inversion of spectral lines and Zeeman-induced Stokes profiles. *Astron. Astrophys.* **577**, A7. (doi:10.1051/0004-6361/201424860)
4. Khomenko E, Collados M. 2015 Oscillations and waves in sunspots. *Living Rev. Sol. Phys.* **12**, 7–21. (doi:10.1007/lrsp-2015-6)
5. Felipe T, Socas-Navarro H, Khomenko E. 2014 Synthetic observations of wave propagation in a sunspot umbra. *Astrophys. J.* **795**, 9. (doi:10.1088/0004-637X/795/1/9)
6. Beckers JM, Tallant PE. 1969 Chromospheric inhomogeneities in sunspot umbrae. *Sol. Phys.* **7**, 351–65. (doi:10.1007/BF00146140)
7. Rouppe van der Voort LHM, Rutten RJ, Sütterlin P, Sloover PJ, Krijger JM. 2003 La Palma observations of umbral flashes. *Astron. Astrophys.* **403**, 277–285. (doi:10.1051/0004-6361:20030237)
8. Centeno R, Collados M, Trujillo Bueno J. 2006 Spectropolarimetric investigation of the propagation of magnetoacoustic waves and shock formation in sunspot atmospheres. *Astrophys. J.* **640**, 1153–1162. (doi:10.1086/500185)
9. de la Cruz Rodríguez J, Rouppe van der Voort L, Socas-Navarro H, van Noort M. 2013 Physical properties of a sunspot chromosphere with umbral flashes. *Astron. Astrophys.* **556**, A115. (doi:10.1051/0004-6361/201321629)
10. Henriques VMJ, Mathioudakis M, Socas-Navarro H. 2017 A hot downflowing model atmosphere for umbral flashes and the physical properties of their dark fibrils. *Astrophys. J.* **845**, 102. (doi:10.3847/1538-4357/aa7ca4)
11. Houston SJ *et al.* 2018 The magnetic response of the solar atmosphere to umbral flashes. *Astrophys. J.* **860**, 28. (doi:10.3847/1538-4357/aab366)
12. Grant SDT *et al.* 2018 Alfvén wave dissipation in the solar chromosphere. *Nat. Phys.* **14**, 480–483. (doi:10.1038/s41567-018-0058-3)
13. Lites BW. 1984 Photoelectric observations of chromospheric sunspot oscillations. II - Propagation characteristics. *Astrophys. J.* **277**, 874. (doi:10.1086/161758)
14. Socas-Navarro H. 2000 Anomalous polarization profiles in sunspots: possible origin of umbral flashes. *Science* **288**, 1396–1398. (doi:10.1126/science.288.5470.1396)
15. Louis RE, Beck C, Ichimoto K. 2014 Small-scale chromospheric jets above a sunspot light bridge. *Astron. Astrophys.* **567**, A96. (doi:10.1051/0004-6361/201423756)
16. Rees DE, Semel MD. 1979 Line formation in an unresolved magnetic element: a test of the centre of gravity method. *Astron. Astrophys.* **74**, 1.
17. Uitenbroek H. 2003 The accuracy of the Center-of-Gravity method for measuring velocity and magnetic field strength in the solar photosphere. *Astrophys. J.* **592**, 1225–1233. (doi:10.1086/375736)
18. Kulander JL, Jefferies JT. 1966 Inference of velocities from line asymmetries. *Astrophys. J.* **146**, 194. IOP Publishing. (doi:10.1086/148868)
19. Keys PH *et al.* 2020 High-resolution spectropolarimetric observations of the temporal evolution of magnetic fields in photospheric bright points. *Astron. Astrophys.* **633**, A60. (doi:10.1051/0004-6361/201936545)
20. Jess DB, Morton RJ, Verth G, Fedun V, Grant SDT, Giagkiozis I. 2015 Multiwavelength studies of MHD waves in the solar chromosphere. *Space Sci. Rev.* **190**, 103–161. (doi:10.1007/s11214-015-0141-3)
21. Stangalini M *et al.* 2018 Propagating spectropolarimetric disturbances in a large sunspot. *Astrophys. J.* **869**, 110. (doi:10.3847/1538-4357/aaec7b)
22. Felipe T, Socas-Navarro H, Przybylski D. 2018 Inversions of synthetic umbral flashes: effects of scanning time on the inferred atmospheres. *Astron. Astrophys.* **614**, A73. EDP Sciences. (doi:10.1051/0004-6361/201732169)

23. Zaghoul MR. 2007 On the calculation of the Voigt line profile: a single proper integral with a damped sine integrand. *Mon. Not. R. Astron. Soc.* **375**, 1043–1048. (doi:10.1111/j.1365-2966.2006.11377.x)
24. Jess DB *et al.* 2019 A chromospheric resonance cavity in a sunspot mapped with seismology. *Nat. Astron.* **4**, 220–227. (doi:10.1038/s41550-019-0945-2)
25. Shine R, Linsky J. 1972 Physical properties of solar chromospheric plages. *Sol. Phys.* **25**, 359–379. (doi:10.1007/BF00192335)
26. Cauzzi G *et al.* 2008 The solar chromosphere at high resolution with IBIS. *Astron. Astrophys.* **480**, 515–526. (doi:10.1051/0004-6361:20078642)
27. de la Cruz Rodríguez J, Socas-Navarro H, Carlsson M, Leenaarts J. 2012 Non-local thermodynamic equilibrium inversions from a 3D magnetohydrodynamic chromospheric model. *Astron. Astrophys.* **543**, A34. (doi:10.1051/0004-6361/201218825)
28. Quintero Noda C *et al.* 2016 Spectropolarimetric capabilities of Ca II 8542 Å line. *Mon. Not. R. Astron. Soc.* **459**, 3363–3376. (doi:10.1093/mnras/stw867)
29. Koons HC, Gorney DJ. 1990 A sunspot maximum prediction using a neural network. *Eos, Trans. Amer. Geophys. Union* **71**, 677. (doi:10.1029/EO071i018p00677-01)
30. Calvo RA, Ceccato HA, Piacentini RD. 1995 Neural network prediction of solar activity. *Astrophys. J.* **444**, 916. (doi:10.1086/175661)
31. Qahwaji R, Colak T. 2007 Automatic short-term solar flare prediction using machine learning and sunspot associations. *Solar Phys.* **241**, 195–211. (doi:10.1007/s11207-006-0272-5)
32. Sudar D, Vršnak B, Dumbović M. 2015 Predicting coronal mass ejections transit times to Earth with neural network. *Mon. Not. R. Astron. Soc.* **456**, 1542–1548. (doi:10.1093/mnras/stv2782)
33. Felipe T, Asensio Ramos A. 2019 Improved detection of far-side solar active regions using deep learning. *Astron. Astrophys.* **632**, A82. (doi:10.1051/0004-6361/201936838)
34. Asensio Ramos A, Requerey IS, Vitas N. 2017 DeepVel: deep learning for the estimation of horizontal velocities at the solar surface. *Astron. Astrophys.* **604**, A11. (doi:10.1051/0004-6361/201730783)
35. Asensio Ramos A, Díaz Baso CJ. 2019 Stokes inversion based on convolutional neural networks. *Astron. Astrophys.* **626**, A102. (doi:10.1051/0004-6361/201935628)
36. Cavallini F. 2006 IBIS: A new post-focus instrument for solar imaging spectroscopy. *Sol. Phys.* **236**, 415–439. (doi:10.1007/s11207-006-0103-8)
37. Schou J *et al.* 2011 Design and ground calibration of the helioseismic and magnetic imager (HMI) Instrument on the Solar Dynamics Observatory (SDO). *Sol. Phys.* **275**, 229–259. (doi:10.1007/s11207-011-9842-2)
38. Pesnell WD, Thompson BJ, Chamberlin PC. 2011 The solar dynamics observatory (SDO). *Sol. Phys.* **275**, 3–15. (doi:10.1007/s11207-011-9841-3)
39. Barnes WT *et al.* 2020 The SunPy Project: open source development and status of the version 1.0 Core Package. *Astrophys. J.* **890**, 68. (doi:10.3847/1538-4357/ab4f7a)
40. Mumford S *et al.* 2020 SunPy: a python package for solar physics. *J. Open Source Softw.* **5**, 1832. (doi:10.21105/joss.01832)
41. Jess DB *et al.* 2012 The source of 3 minute magnetoacoustic oscillations in coronal fans. *Astrophys. J.* **757**, 160. (doi:10.1088/0004-637X/757/2/160)
42. Pietarila A, Socas-Navarro H, Bogdan T. 2007 Spectropolarimetric observations of the Ca II λ 8498 and λ 8542 in the Quiet Sun. *Astrophys. J.* **663**, 1386–405. (doi:10.1086/518714)
43. Pietarila A, Socas-Navarro H, Bogdan T. 2007 Spectropolarimetric inversions of the Ca II 8498 and 8542 Å lines in the quiet sun. *Astrophys. J.* **670**, 885–902. (doi:10.1086/521271)
44. de la Cruz Rodríguez J, De Pontieu B, Carlsson M, Rouppe van der Voort LHM. 2013 Heating of the magnetic chromosphere: observational constraints from Ca II λ 8542 spectra. *Astrophys. J.* **764**, L11. (doi:10.1088/2041-8205/764/1/L11)
45. Virtanen P *et al.* 2020 SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272. (doi:10.1038/s41592-019-0686-2)
46. Virtanen P *et al.* 2019 scipy/scipy: SciPy 1.4.1. Zenodo.
47. McLean AB, Mitchell CEJ, Swanston DM. 1994 Implementation of an efficient analytical approximation to the Voigt function for photoemission lineshape analysis. *J. Electron. Spectrosc. Relat. Phenom.* **69**, 125–132. (doi:10.1016/0368-2048(94)02189-7)
48. Pedregosa F *et al.* 2011 Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* **12**, 2825–2830.

49. Buitinck L *et al.* 2013 API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD workshop: Languages for data mining and machine learning* (eds B Crémilleux, L De Raedt, P Frasconi, T Guns), pp. 108–122. ECML PKDD.

50. Grisel O *et al.* 2020 scikit-learn/scikit-learn: Scikit-learn 0.22.1. Zenodo.

51. Costa M, Filippi E, Pasero E. 1994 Multi-layer perceptron ensembles for pattern recognition: some experiments. In *Proc. of 1994 IEEE Int. Conf. on Neural Networks (ICNN'94), Orlando, FL, 28 June–2 July 1994*. New York, NY: IEEE.

52. Byrd RH, Peihuang L, Nocedal J. 1996 A limited-memory algorithm for bound-constrained optimization. Oak Ridge, TN: Office of Scientific and Technical Information (OSTI). March. (doi:10.2172/204262)

53. Hahnloser RHR, Sarpeshkar R, Mahowald MA, Douglas RJ, Seung HS. 2000 Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature* **405**, 947–951. (doi:10.1038/35016072)

54. Bergstra J *et al.* 2010 Theano: A CPU and GPU math compiler in python. In *Proc. of the 9th Python in Science Conf., Austin, TX, June 28–July 3 2010*, pp. 3–10. SciPy.

55. Kentischer TJ *et al.* 2012 The visible tunable filtergraph for the ATST. In *Ground-based and Airborne Instrumentation for Astronomy IV* (eds McLean IS, Ramsay SK, Takami H), pp. 2156–2171. Bellingham, WA: SPIE.

56. Kingma DP, Ba J. 2014 Adam: a method for stochastic optimization. (<http://arxiv.org/abs/1412.6980>).

57. Branch MA, Coleman TF, Li Y. 1999 A subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained minimization problems. *SIAM J. Sci. Comput.* **21**, 1–23. (doi:10.1137/S1064827595289108)

58. Kuckein C *et al.* 2017 Ca II 8542 Å brightenings induced by a solar microflare. *Astron. Astrophys.* **608**, A117. (doi:10.1051/0004-6361/201731319)

59. Samanta T *et al.* 2016 The effects of transients on photospheric and chromospheric power distributions. *Astrophys. J.* **828**, 23. (doi:10.3847/0004-637X/828/1/23)

60. Löhner-Böttcher J, Bello González N. 2015 Signatures of running penumbral waves in sunspot photospheres. *Astron. Astrophys.* **580**, A53. (doi:10.1051/0004-6361/201526230)

61. Houston SJ *et al.* 2020 Magnetohydrodynamic nonlinearities in sunspot atmospheres: chromospheric detections of intermediate shocks. *Astrophys. J.* **892**, 49. (doi:10.3847/1538-4357/ab7a90)

62. Tian H *et al.* 2014 Prevalence of small-scale jets from the networks of the solar transition region and chromosphere. *Science* **346**, 1255711–1255711. (doi:10.1126/science.1255711)

63. Kuridze D *et al.* 2015 H α line profile asymmetries and the chromospheric flare velocity field. *Astrophys. J.* **813**, 125. (doi:10.1088/0004-637X/813/2/125)

64. Sharma R, Verth G, Erdélyi R. 2017 Dynamic behavior of spicules inferred from perpendicular velocity components. *Astrophys. J.* **840**, 96. (doi:10.3847/1538-4357/aa6d57)

65. Dieleman S, Willett KW, Dambre J. 2015 Rotation-invariant convolutional neural networks for galaxy morphology prediction. *Mon. Not. R. Astron. Soc.* **450**, 1441–1459. (doi:10.1093/mnras/stv632)

66. Wright DE *et al.* 2015 Machine learning for transient discovery in Pan-STARRS1 difference imaging. *Mon. Not. R. Astron. Soc.* **449**, 451–466. Oxford University Press (OUP). (doi:10.1093/mnras/stv292)

67. Kuckein C, González Manrique SJ, Kleint L, Asensio Ramos A. 2020 Determining the dynamics and magnetic fields in the chromospheric He I 10830 Å triplet during a solar filament eruption. *Astron. Astrophys.* **640**, A71. (doi:10.1051/0004-6361/202038408)

68. De Pontieu B, Erdélyi R, James SP. 2004 Solar chromospheric spicules from the leakage of photospheric oscillations and flows. *Nature* **430**, 536–539. (doi:10.1038/nature02749)

69. Socas-Navarro H, Pillet VM, Elmore D, Pietarila A, Lites BW, Sainz RM. 2006 Spectropolarimetric observations and non-Lte modeling of ellerman bombs. *Sol. Phys.* **235**, 75–86. (doi:10.1007/s11207-006-0049-x)

70. Rezaei R, Beck C. 2015 Multiwavelength spectropolarimetric observations of an Ellerman bomb. *Astron. Astrophys.* **582**, A104. (doi:10.1051/0004-6361/201526124)

71. Schmidt W *et al.* 2016 End-to-end simulations of the visible tunable filter for the Daniel K. Inouye Solar Telescope. In *Ground-based and airborne instrumentation for astronomy VI* (eds CJ Evans, L Simard, H Takami). Bellingham, WA: SPIE.

72. Tritschler A *et al.* 2016 Daniel K. Inouye Solar Telescope: High-resolution observing of the dynamic Sun. *Astronomische Nachrichten* **337**, 1064–1069. (doi:10.1002/asna.201612434)
73. Beck C, Choudhary DP, Rezaei R, Louis RE. 2014 Fast Inversion of solar Ca II spectra. *Astrophys. J.* **798**, 100. (doi:10.1088/0004-637X/798/2/100)
74. MacBride CD, Jess DB. 2020 MCALF v0.1. Zenodo. (doi:10.5281/zenodo.3924527)