University of Sheffield

Neural Network Based Flow Recovery and Magnetic Structure Identification in the Solar Atmosphere

Matthew George Lennard

Supervisors: Professor Viktor Fedun & Doctor Gary Verth

A report submitted in partial fulfilment of the requirements
for the degree of doctor of philosophy in solar and space plasma physics

in the

School of Electrical and Electronic Engineering

August 26, 2025

Declaration

All sentences or passages quoted in this document from other people's work have been specifically acknowledged by clear cross-referencing to author, work and page(s). Any illustrations that are not the work of the author of this report have been used with the explicit permission of the originator and are specifically acknowledged. I understand that failure to do this amounts to plagiarism and will be considered grounds for failure.

Name: Matthew George Lennard

Signature:

Date: 26/08/2025

Acknowledgements

First and foremost I express my gratitude to the Science and Technology Facilities Council (STFC) UK for funding the research which is presented in this thesis.

This Ph.D. has been an immense personal journey for myself and I have faced many bumps in the road along the way to completing this work. However, the immense positives gained from meeting new people, exploring new places and discovering new knowledge has made everything worthwhile.

For the endless support of my parents and continuing encouragement from my brother, I am eternally grateful. Additionally, I want to thank my dear friends who have constantly provided their love and support during the last 4 years. Attempting to name them all would be a disservice to anyone who may be forgotten but I could not have made it here without any of them. Especially, Slava.

I thank those who have sat alongside me during these last four years: Samuel, Jack, Max, Farhad, and Ellie. I wish them all the best and I am nothing but glad to have met them because of this opportunity.

I give my gratitude to Andrés, Benoit and Suzana, who have been excellent mentors, and thank my examiners, Dave and Leon, for their support in refining the matter of this thesis.

Thank you to my supervisors for their guidance and support these past few years, especially Professor Viktor Fedun whose kindness and patience has delivered me to where I am today. Thank you for believing in me.

Lastly, I'd like to give a final thanks to myself for not giving up. This Ph.D. has been my greatest test and it could not have been possible if not for the great people supporting me.

Summary

Energy transport is a fundamental aspect of solar atmospheric dynamics. Plasma motions can generate concentrated magnetic flux, driving large-scale transients such as solar flares and coronal mass ejections, which pose an increasing threat to space- and ground-based infrastructure. However, direct measurement of plasma flows on the Sun remains unfeasible, raising a critical question: how can plasma flows be accurately recovered from observations?

In a step toward answering this question this thesis employs the neural network DeepVel for estimating flows in new scenarios including active regions and the chromosphere, thus enabling the estimation of flows by assimilating data from state-of-the-art simulations. DeepVel was further extended to higher regions of the solar atmosphere, where a reduced optical thickness results in less coherent apparent motions. By training on a simulation containing swirling motions in the chromosphere, the networks ability to identify coherent swirls against an incoherent background was tested using a vortex detection method. Network performance was evaluated by undergoing cross-validation with simulations, performing error analysis and comparing with the widely-used Fourier-based local correlation tracking technique. Its ability to imitate the physics present was also tested for the first time by identifying through Lagrangian coherent structures in the recovered flows.

Results highlight that DeepVel is highly capable for identifying coherent flow structures, that determine the evolving flow dynamics, across simulations. These flow structures were shown to correspond well to the presence of emerging active regions thus presenting a promising performance for use with real-world observations for forecasting and tracking the presence of active regions. Despite currently available data, both simulated and observed, tests indicate success in applying DeepVel to chromospheric plasmas.

Contents

1 Introduction **1**

1.1 Solar Structure 2

1.2 Observations of the Sun 12

1.3 Magnetohydrodynamics in the Sun 13

1.3.1 MHD Waves 17

1.3.2 MHD Flows 22

1.3.3 Modelling the Sun 24

1.4 Flow Recovery methods 25

1.4.1 Traditional Flow Recovery 25

1.4.2 Neural Networks 28

1.4.3 DeepVel 30

1.5 Lagrangian Flow Analysis 32

1.5.1 Coherent Structures and Lagrangian Coherent Structures 33

1.5.2 Finite-Time Lyupanov Exponent 33

1.6 Thesis outline 35

2 Comparing DeepVel and Fourier-Local Correlation Tracking for Flow Recovery in Simulated Active Regions **38**

2.1 Introduction 38

2.2 Methods 41

2.2.1 Data 41

2.2.2 Preparing FLCT 44
2.2.3 Training DV 47
2.2.4 Testing 51
2.3 Results 53
2.3.1 Metrics 53
2.3.2 Comparing FLCT and DV performance with MURaM 54
2.4 Summary 66

3 Recovering Lagrangian Coherent Structures with DeepVel 69
3.1 Introduction 69
3.2 Methods 72
3.3 Results 73
3.3.1 Coherent Structures in the R2D2 Simulation 74
3.3.2 Coherent Structures in the MURaM Simulation 77
3.4 Summary 79

4 Identifying and Tracking the Presence of Active Regions with Coherent Photospheric Flow Structures 84
4.1 Introduction 84
4.2 Methodology 88
4.2.1 The Finite-Time Lyapunov Exponent 88
4.2.2 Identifying Magnetic structures 89
4.3 Results 89
4.4 Summary 107

5 Identifying Coherent Swirling Structures in the Chromosphere 109
5.1 Introduction 109
5.2 Methods 113
5.2.1 Data 113
5.2.2 DeepVel 115

5.2.3 Vortex Detection Methods 116
5.3 Results 117
5.3.1 DeepVel-heights 117
5.3.2 DeepVel-chromosphere 119
5.3.3 Validation with the λ_2-Criterion 123
5.4 Summary 124

6 Discussion and Conclusions **128**

7 Future Works **132**

A Point Spread Functions **135**

B P-modes **140**

List of Figures

1.1 Diagram from Priest (2014), showing a rough description of the internal and external structure of the Sun. At the centre is the dense core in which nuclear fusion takes place. Energy radiates outward from the core in the radiative zone which heats and drives the convective region. The surface of this region is the first visible layer of the atmosphere, the photosphere. The photosphere transitions into the chromosphere and then into the corona which fills the entire heliosphere. 4

1.2 The above figure depicts the structure of the solar atmosphere based on height, temperature and density. Between 0 and 1000km above the surface lies the first visible layer, the photosphere. The photosphere is the densest region of the atmosphere and is also home to the coolest region of the solar exterior. The next region, by height, is the chromosphere, this region is visible in strong spectral lines and is significantly less dense than the photosphere, where the magnetic field begins to dominate. In the upper chromosphere lies the transition region where there is a sudden heating of the atmosphere and a rapid decline in the density. The corona is the uppermost region of the solar atmosphere. It fills out the entire heliosphere, which encloses the solar system. The corona has a temperature range in the millions of Kelvin, several Mm from the solar surface and is sparse in plasma, but is dominated by magnetic fields, i.e. $\beta < 1$. The figure was taken from Lang (2007). 5

1.3 Examples of photospheric images from the helioseismic and magnetic imager (HMI) instrument, onboard the solar dynamics observatory (SDO). Shown on the left are magnetogram images, where dark regions indicate negative polarity of the magnetic field and the lighter regions have a positive polarity. In the centre, the continuum intensity is shown, where the dark regions indicate a drop in intensity. On the right, the Doppler measurements over the disk of the Sun are shown. The lighter regions indicate motion away from away from the viewer; note here that the rotation of the Sun directly influences measurements. Images courtesy of NASA. 6

1.4 Figure depicting some of the key Fraunhofer lines and the elements responsible for their formation. Each sharp dip in the spectral irradiance highlights a distinct spectral line. Here, the H-α, -β, Fe and Ca lines, are of specific interest due to their correspondence with the solar photosphere and chromosphere. Namely, the H-α and -β lines and the Ca lines are used to observe the chromosphere and the Fe lines are used to study the photosphere. Image taken from Cyamahat (2022) 7

1.5 Example of an image of the chromosphere, highlighting a brightening caused by footpoints of a magnetic loop structure, taken in the Ca II (K line of the Fraunhofer spectrum) taken by SST/CHROMIS. Image taken from Testa et al. (2023). 8

1.6 Figure showing the Vernazza, Avrett and Loeser (VAL) model for the solar atmosphere. The axes show the temperature of regions of the solar atmosphere against their respective heights. This highlights the formation heights of spectral lines present in observations. The figure was taken from Vernazza et al. (1981). 9

1.7 A map of the range of values that the plasma-β may take as a function of height in the solar atmosphere, taken from Gary (2001). The heights shown here are approximate as the location of each region is determined by a number of factors. 21

1.8 Sketch of the architecture of the DeepVel neural network from Asensio Ramos, A. et al. (2017) showing the layers and transformations acting on an input (left) to the output (right). The highlighted block reveals the combination of layers that form the residual blocks of the network. 31

2.1 An example of the fully developed AR at the time of peak magnetic flux (i.e. 60 hours after the start of simulation), from the R2D2 magnetoconvection simulation. On the top row, two panels show the white light intensity and the bottom row shows the corresponding magnetograms. The right-hand side shows a close-up view of the intensity and magnetogram of the region bounded by the green square on the left. The intensity was normalised by the maximum intensity of the entire region. 42

2.2 An example slice of the AR at 83 hours after the start of the simulation, from the MURaM simulation. The panels in the left-hand column show the continuum intensity (a) and the magnetogram (b). The right-hand side panels show a close-up view of the intensity (top panel) and magnetogram (bottom panel) of the region bounded by the green square. The intensity was normalised between to be on a scale between 0 and 1. 43

2.3 The distribution of normalised intensity from the R2D2 (red) and MURaM (blue) simulations. 44

2.4 Comparison of the synthesised continuum intensity from the R2D2 simulation with the same image with a Gaussian blur applied, typically used for the application of LCT methods. The value of σ in corresponds to a FWHM of $\approx$ 1200km, as was found to be the best sampling size by Verma et al. (2013). Note that only granular- and above-sized structures remain after the sampling. 45

2.5 Velocity fields from (a) simulation, (b) the velocity field recovered from simulation using FLCT on 2 consecutive frames and (c) the average of 15 recovered velocity fields. The background is coloured by the Gaussian-smoothed (FWHM = 1200km) continuum intensity of the simulation. 47

2.6 Scatter plots showing how time-averaging of velocity fields produced by FLCT improves correlation between the normalised recovered and true speed of plasma. The curves of best fit for each distribution is shown in red and the ideal case is shown in black. Panel (a) shows the result of recovering only a single frame. Panel (b) show the 30 minute-averaged speeds. 48

2.7 Loss during the first 100 epochs of training from the DV during training. The black curve represents the loss value from the training data at each epoch and blue curve represents the loss value for the validation set (calculated at the end of each epoch). The red-dotted curve shows the minimum validation loss for the network at each epoch and the green point shows the minimum validation loss over the whole training process. 50

2.8 A close view of the divergence of the horizontal components of velocity, 2 hours into the MURaM simulation. This figure highlights a region where > kG magnetic flux emerges later in the simulation. Panel (a) shows the divergence field from simulation over a zoomed-in section of the region highlighted in figure 2.2. Panel (d) shows a closeup of the LIC with the velocity field superimposed on the divergence field of the highlighted region from panel (a). Panels (b) and (c) show the same regions as panels (a), where the velocities have been estimated by FLCT and DV, respectively. Panels (e) and (f) show the same region as (a), again using velocities from FLCT and DV, respectively. 55

2.9 The same visualization as Figure 2.8 is shown, at a time after the AR has emerged (83 hours). The zoomed-in regions highlight the boundary of a pore, containing complex flow structures such as saddles and vortices. 57

2.10 Histograms showing the distribution of different components of the velocity field at a time of 83 hours, for the original simulation (solid coloured backgrounds) and the flows recovered by FLCT (the outlined curves) over regions of $> kG$ magnetic flux (blue), $< kG$ flux (magenta) and over the entire field of view (green). Panels (a) and (b) show the distributions for the x and y components of velocity. Panel (c) shows the distribution of speed and (d) shows the distribution of divergence. All are shown in cgs units. Note the stark contrast in shapes of the distributions from FLCT compared to the simulation which indicate the method is unable to replicate the target distribution which will lead to large errors in the velocity fields. 58

2.11 Same as fig. 2.10 except using recovered velocities from $\mathrm{DV}_{\mathrm{R2D2}}$. Note here the much closer shapes in distributions compared to the previous figure, indicating that the performance from DV is a significant improvement over FLCT. . . . 59

2.12 The cumulative distribution function for the horizontal velocity components (a) v_x and (b) v_y, and the (c) divergence of the 2D velocity field $\nabla \cdot \mathbf{v}$. The CDF for the MURaM velocities are shown as the filled-in grey background, and the CDFs for DV and FLCT are shown in blue and orange, respectively. Note that the CDFs from DV only differs a small amount from MURaM's, except in divergence where the gradient is relatively steeper, whereas FLCT produces CDFs that are significantly steeper and differ a lot from the target distribution. 60

2.13 Scatter plots of components of velocities (a) and (b) and their divergence (c) recovered from by FLCT and their lines of best fit, plotted against the simulated values from MURaM, shown in cgs units. Values over ARs are shown in blue, QS are shown in purple. The green line shows the overall line of best fit over all types of regions and the black line shows the target line of best fit ($y = x$). Points have had their opacity lowered to remove outliers and highlight where the strongest spread of velocities lie. 60

2.14 Same as fig. 2.14, however using comparing velocities predicted by DV_{R2D2} against the simulation velocities. 61

3.1 An observation of Jupiter's atmosphere, namely the great red spot, from the Voyager 1 mission in February 1979. Image courtesy of NASA/JPL and Björn Jónsson. There is a distinct stratification of the gas, separated by material surfaces which distinguish Lagrangian coherent structures such as an elliptical LCS surrounding the spot, acting as a transport barrier. 71

3.2 The normalised forward- (red) and backward-FTLE (blue) field superimposed on the magnetic field, corresponds to the highlighted region in fig. 2.1 in the R2D2 simulation at the time of peak magnetic flux in the photosphere (60 hours). Panels (a) and (e) show the magnetic field averaged over the time of integration. (b) and (f) show the backward-FTLE ridges resulting from a 20-minute integration of the simulated and recovered velocity fields, respectively. Panels (c) and (g) show the resulting forward integrated FTLE fields. Panels (d) and (h) both the forward and backward integrated FTLE fields superimposed on the magnetic field. The ridges resulting from the 20-minute integration correspond to features on the length scales of granules. 75

3.3 The same as Figure 3.2 however the FTLE ridges here result from a 100-minute integration corresponding to mesogranular features. 76

3.4 Distributions of the 20-minute forward-FTLE field, at various times in the R2D2 simulation during the AR lifetime: 1h 40min (early, when there is little concentration of magnetic flux in the photosphere; dark green), 40h (during the emergence of the AR in the photopshere; blue), 60h (the time of peak magnetic flux density in the photopshere; magenta), 166h 40min (close to the time of total decay of the AR; yellow). The top row shows the FTLE distributions calculated from the simulation velocities and the bottom shows the distributions calculated using velocities predicted by DV_{R2D2}. The left column shows the distributions over a region which contains $<$ kG magnetic flux density throughout the evolution of the photospheric surface and the right column displays the distributions over a region where $>$ kG flux is present during the emergence and peak of the AR. The vertical lines show the corresponding positive foward-FTLE means for each distribution 81

3.5 The same as fig. 3.2, however highlighting the region from the MURaM simulation shown in fig. 2.2. 82

3.6 The same as fig. 3.3, however highlighting the region from the MURaM simulation shown in fig. 2.2. 82

3.7 Distribution of the 20-minute forward-FTLE field at times of 2h (prior to the emergence of the AR; blue) and 83h (post-emergence of the AR; magenta) in the MURaM simulation. The top row shows the FTLE distributions calculated using velocities from the simulation and the bottom rows shows the distributions using velocities recovered by $\mathrm{DV}_{\mathrm{R2D2}}$. The left column shows the distributions over a region, which contains $<$ kG magnetic flux density throughout the simulation and the right column shows the distributions over a region where $>$ kG magnetic flux density is present when the AR emerges. The vertical lines highlight the respective positive forward-FTLE distribution means. 83

4.1 Plot from Silva et al. (2023), showing the normalised distance between peaks in the 100-minute forward FTLE field in the R2D2 simulation around the time of emergence, which was chosen to be t_0 by the authors. The blue line corresponds to the average distance between peaks (P_d^*) of the forward FTLE ridges and the orange line shows the evolution of the maximum unsigned flux ($B_{\max}$) in Gauss. The green shaded region shows when a significant enough change in the distances between peaks to identify the presence of an emerging AR 86

4.2 Plot from Silva et al. (2023) showing the time-evolution of the magnetic field at the $\tau = 1$ surface from the R2D2 simulation about the time of AR emergence in the photosphere, t_0. 87

4.3 Plot from Silva et al. (2023) showing the time-evolution of the 100-minute forward FTLE field over the photosphere about the time of AR emergence, t_0. Distinct material lines form over the region where the magnetic flux is strongest up to 10 hours before the full emergence. 87

4.4 The 20-minute forward-FTLE fields from R2D2 simulation velocities (left column), magnetic flux (middle) and confusion matrices of the regions classified as active via the FTLE field vs the magnetic field (right). The top, second, third and fourth rows display key times throughout the sunspot evolution of 3 (mostly quiet), 40 (emerging sunspot), 60 (peak photospheric magnetic flux) and 170 (near full decay of sunspot) hours. Cases identified as active via the FTLE field are highlighted in pink when the FTLE mean value differs by 0.25 from the mean FTLE value in the QS. The actual positive cases, determined by the concentration of magnetic flux, are highlighted in blue. 104

4.5 The same as fig. 4.4, however the velocities used for calculating the FTLEs were estimated by DV. 105

4.6 ROC curves for 20 evenly spaced thresholds of the relative positive forward-FTLE mean difference and 4 different spatial scales (depicting the length of one side of the tile). Each point along the curve shows the FPR and TPR for each threshold. Each curve is labelled with the area under the curve (AUC). The diagonal, dashed line represents the possible case and the dotted line shows the best case scenario, where the optimum case is highlighted by the black point. Panel (a) shows the curves when using the velocities from the R2D2 simulation and (b) shows the curves when using velocities estimated by DV for calculating the FTLEs. 106

5.1 Image taken from Wedemeyer-Böhm, S. & Rouppe van der Voort, L. (2009) showing the swirling motions in the chromosphere with corresponding photospheric bright points. The top 3 rows show images taken in the SST CRISP wide-band (photosphere), Ca II IRT 8452Å line wing (photosphere), and line core (chromosphere) respectively. The bottom row shows Doppler measurements. 110

5.2 Magnitude of the horizontal velocities from the Bifrost simulation 'ch024031'. These are shown at heights of 0Mm, 1Mm and 2Mm relative to the surface of the Sun, displaying the photosphere, mid-chromosphere and upper-chromosphere. 114

5.3 Example of (a) an image synthesised in 3mm band of the ALMA radio telescope, from the Bifrost simulation (ch024031), (b) the radiation formation heights corresponding to the image and (c) the magnitude of the horizontal velocity at the formation heights. 116

5.4 The (a) radiation formation heights of a frame of the Bifrost simulation highlighted in fig. 5.3, (b) heights recovered by the version of DV trained to recover heights from 3mm band images of the chromosphere and (c) the relative error of the heights produced by DV. 117

5.5 The (a) heights over the field of view of a frame from the Bifrost simulation. The highlighted regions depicts a slice containing swirling features, where the averaged true heights (blue curve) from simulation and the predicted heights by $\mathrm{DV_h}$ (orange curve) are compared on the left-hand axis of panel (b), along with the relative error of the predicted heights (dashed line) on the right-hand axis. Panel (c) depicts a histogram of the height distribution from simulation (blue) and predicted by $\mathrm{DV_h}$ (orange) across the entire field of view. 118

5.6 The (a) 3mm intensity of simulated chromosphere from Bifrost, normalised by the mean intensity, (b) the true values of the horizontal speed at the heights of the 3mm band line formation, (c) the heights of the 3mm band line formation, (d) the heights predicted by DV of the 3mm band line formation (see fig. 5.4), (e) the velocities predicted by DV using both the true heights and (f) the predicted heights 120

5.7 The speed (shown by the colouration) and orientation of the same swirling structure from (a) the Bifrost simulation, (b) recovered by $\mathrm{DV_h}$ using a synthesised 3mm band image and the radiation formation heights from simulation, and (c) the same as (b) but using heights predicted by DV_h. 123

5.8 Vortex cores detected in the simulation flows as well as in flows recovered by DV using both the true formation heights and predicted formation heights superposed on the 3mm intensities. 124

A.1 This figure depicts (a) Original synthesised image from simulation, (b) log plot of PSF to show the effect of the lens and (c) the resulting image synthesised to look like one from the IMAX telescope after convolving with its PSF. . . . 138

A.2 This figure depicts the divergence field over a part of the MURaM simulation (b) and the LIC visualisation of the flows, superposed on the divergence in a region with < kG magnetic flux (a), and a region with > kG magnetic flux (c). The same region as (b) is shown for velocity fields produced by DV when using the images convolved with the IMAX PSF in (e). The same < kG and > kG regions, described above, are also shown in (d) and (f), using the DV-recovered velocities from images convolved with the IMAX PSF. 139

B.1 Variation of continuum intensity over time against the Pearson correlation of horizontal velocities recovered by DV with the true horizontal velocities from the Stagger simulation. The intensity varies over time due to 5-minute oscillations induced by p-modes. The recovered velocities show negligible variation in their high correlation with the true velocities. 141

B.2 Example flows from the Stagger simulation (left column) versus the flows predicted by $\mathrm{DV_{R2D2}}$ (right column). The regions are coloured by the horizontal divergence field and the bottom highlight a region where a LIC field has been superposed on the divergence to highlight the velocity field lines. 142

B.3 The 20-minute backward- (left) and forward- (right) FTLE fields calculated using the flows from the Stagger simulation (top row) and velocities produced by $\mathrm{DV_{R2D2}}$ (bottom row). 143

List of Tables

2.1 Table displaying the quantitative properties of the distributions of the horizontal velocity fields ($\mathbf{v}_h$) from the R2D2 simulation and those recovered from $\mathrm{DV}_{\mathrm{R2D2}}$. 64

2.2 Table displaying quantitive properties of the distributions of the horizontal velocity fields ($\mathbf{v}_h$) from the MURaM simulation and those recovered from FLCT and $\mathrm{DV}_{\mathrm{R2D2}}$. 64

2.3 Metrics for the horizontal velocity fields recovered by $\mathrm{DV}_{\mathrm{R2D2}}$ from the testing dataset of the R2D2 simulation. The values are split over QS conditions ($<$ kG) and AR conditions ($>$ kG). The RMSE for velocities and derivatives is given in units of kms^{-1} and s^{-1}, respectively. 65

2.4 Metrics comparing the recovered horizontal velocity field from FLCT and $\mathrm{DV}_{\mathrm{R2D2}}$ to the original velocities from the MURaM simulation. Velocities were taken over the entire $\tau = 1$ surface at a time where the $>$ kG magnetic flux was present in the photosphere (i.e 83 hours after the beginning of the simulation). Compared values are split over QS conditions ($<$ kG) and AR conditions ($>$ kG). The RMSE for velocities and derivatives is given in units of kms^{-1} and s^{-1}, respectively. 65

4.1 Performance of various relative FTLE mean differences, calculated from the R2D2 simulation (Hotta & Iijima 2020) velocities over various grid spacings as a classifier for AR levels of magnetic flux. 91

4.2 Performance of various relative FTLE mean differences, calculated using velocities recovered by DV from the R2D2 simulation (Hotta & Iijima 2020) intensities, over various grid spacings as a classifier for AR levels of magnetic flux. 96

5.1 Metrics showing the performance of the DV_h and DV_c networks. DV_c was tested in two cases, one where the true radiation formation heights (TH) were used for inputs and one where the predicted heights (PH) from DV_h were used. 122

Chapter 1

Introduction

Our nearest star, the Sun, is responsible for providing energy to, and maintaining of, all life on Earth. It is home to a great number of phenomena, driven by its magnetic field, which are constantly interacting with the magnetosphere of the Earth. The effects of this interaction is called space weather (SW; see Schwenn 2006, for details of its nature and impact). Examples of phenomena which strongly influence SW are solar flares (SFs) and coronal mass ejections (CMEs). These are a result of the reconnecting of magnetic fields in the solar atmosphere, have the potential of disturbing the Earth's magnetic field and introducing an influx of highly energetic ions into the ionosphere. These disturbances and particles have the potential to cause damage to space- and, in extreme cases, ground-based infrastructure. Perhaps the most famous recent example is when a CME caused an atmospheric disturbance on Earth, causing a rocket, carrying Starlink satellites into orbit, to burn up and re-enter the atmosphere. In the case of ground-based infrastructure, in 1989 a powerful SF caused a large perturbation in the magnetosphere, causing a build-up of currents at the Earth's surface, resulting in disrupting a power station in Quebec and thus causing a 12 hour power outage (Boteler 2019). Therefore, understanding the Sun in its entirety is of great interest, not only for the curiosity of mankind but also for protecting the infrastructure that is necessary for maintaining our technology-dependent lifestyle. The transport of energy, including the magnetic field, across the solar surface and between layers of the atmosphere is largely driven by fluid motions of plasma, of

which the Sun is mostly comprised. To understand how plasma flows influence the magnetic field topology of the Sun, we need to be able to identify such systems and produce effective physical models from real-world data. The Sun provides the largest natural laboratory for magnetised plasma and thus should be studied extensively. However, a key problem with studying motions in the solar atmosphere is that the velocity field of the plasma cannot be measured directly, and only inferred through other means. Thus a hugely important question in the literature for understanding the dynamics of the Sun and processing the large swathes of data being collected is: how can plasma flows be inferred from observations of the solar atmosphere, both accurately and efficiently? This thesis aims to answer this by testing a neural network approach for assimilating data from simulations of the solar atmosphere for application with observations.

1.1 Solar Structure

The Sun is highly stratified, with each stacked layer being distinguished from the others by its density, temperature and dynamics (see fig. 1.1 for more details). Deep within the core of the Sun, energy is produced by means of nuclear fusion, this energy is driven outwards by different mechanisms, including radiative diffusion and convection. At the surface of the Sun, the atmosphere begins, where plasma moves more freely and energy can be transported along magnetic fields. The solar atmosphere is dubbed the heliosphere, which encompasses the entire solar system (see Priest 2014). In the following paragraphs, these layers are described in more detail.

Solar Interior

The interior of the Sun has been determined through helioseismology (e.g., see Gizon & Birch 2005), similar to seismology on Earth, different dynamics are determined by measuring how waves propagate through the solar interior.

Starting from the centre, the core of the Sun is an incredibly dense region, with density of $\approx 1.54 \times 10^5 \mathrm{kgm}^{-3}$, in which nuclear fusion takes place as a result of extreme temperatures

and confinement due to gravity, at the surface $g_\odot = 274\text{ms}^{-2}$ ($\approx 28\times$ Earth gravity). The nuclear fusion of mainly hydrogen into helium produces kinetic energy and thus exerts an outward pressure, preventing the collapse of the star. The core accounts for $0.25\text{R}_\odot$, where $1\text{R}_\odot = 695.5\text{Mm}$ is the solar radius (Priest 2014).

Radiative zone

Beyond the core is the largest part (0.25–$0.7\text{R}_\odot$) of the solar interior, the radiative zone. In this region, energy is transported out from the core by means of radiative diffusion.

Convective region

The convective region (0.7–$1\text{R}_\odot$), which sits on top of the radiative zone, transports heat mainly through means of convection. Here, plasma is heated by radiation at the bottom and then, under buoyancy, rises to the solar surface, known as the photosphere (the first visible layer), where the plasma cools via radiation, causing the plasma to fall and repeat the process. It is presumed that within this region, a dynamo action takes place in which electrically charged plasma generates magnetic fields through shearing motions, however, the exact nature of how the Sun's magnetic field is generated is uncertain.

The plasma in the solar interior is so incredibly dense that if one were to trace the path of energy produced in the core, the photons responsible for transporting this energy take hundreds of thousands of years to reach the surface due to the photons being reabsorbed and re-emitted. Whereas if the photon was unimpeded and travelling through a vacuum, it would cover the same distance in a little over 2 seconds (Priest 2014).

Solar Atmosphere

The solar atmosphere is highly stratified and contains a wide range of plasma dynamics including sources for wave excitation. It consists of several layers that are split into the photosphere, the chromosphere and the solar corona. The aforementioned layers form the outer region of the Sun which is optically translucent and hence visible in a range of spectra. The

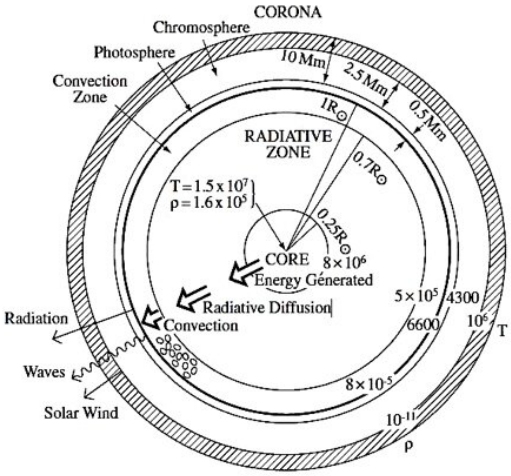

Figure 1.1: *Diagram from Priest (2014), showing a rough description of the internal and external structure of the Sun. At the centre is the dense core in which nuclear fusion takes place. Energy radiates outward from the core in the radiative zone which heats and drives the convective region. The surface of this region is the first visible layer of the atmosphere, the photosphere. The photosphere transitions into the chromosphere and then into the corona which fills the entire heliosphere.*

regions are split into regions based on dynamics and properties such as density, temperature and optical thickness.

The optical thickness of a medium defines how transparent it is. Suppose the intensity of radiation at the source of an object is given by I_0. Then the observed intensity I is measured to be

$$I = I_0 e^{-\tau}, \tag{1.1}$$

where we call τ the optical thickness (Priest 2014). That is, the optical thickness defines how much radiation remains not scattered or unabsorbed by a medium at the point of observation.

The first visible layer of the Sun is the photosphere, which is often labelled as the surface. It measures up to a few hundred km above the top of the convective region. This region has optical thickness $\tau \lesssim 1$ in near-ultraviolet, visible and near-infrared light, however it is optically thick in most strong spectral lines. Because of the optical thickness across the light spectrum, the photosphere is typically observed with white light or continuum intensity which measures radiative intensity over all wavelengths. The entire photosphere facing the Earth

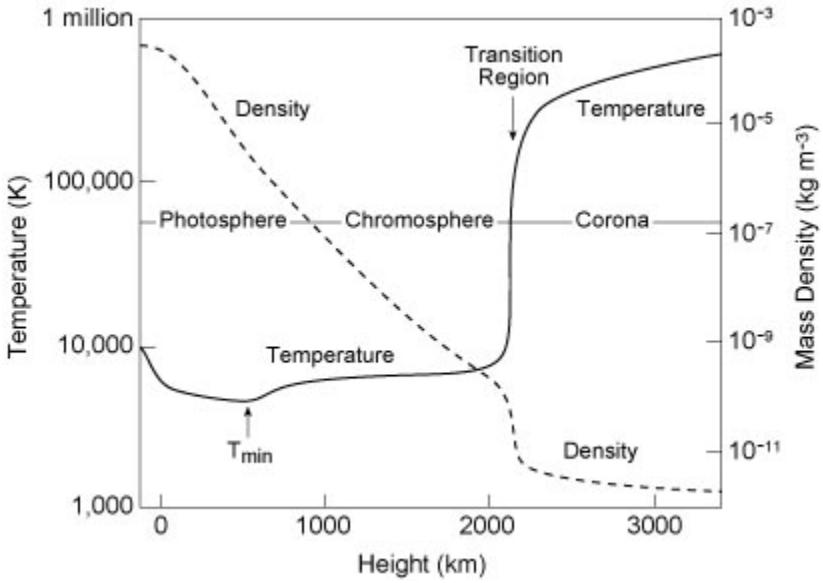

Figure 1.2: *The above figure depicts the structure of the solar atmosphere based on height, temperature and density. Between* 0 *and* 1000km *above the surface lies the first visible layer, the photosphere. The photosphere is the densest region of the atmosphere and is also home to the coolest region of the solar exterior. The next region, by height, is the chromosphere, this region is visible in strong spectral lines and is significantly less dense than the photosphere, where the magnetic field begins to dominate. In the upper chromosphere lies the transition region where there is a sudden heating of the atmosphere and a rapid decline in the density. The corona is the uppermost region of the solar atmosphere. It fills out the entire heliosphere, which encloses the solar system. The corona has a temperature range in the millions of Kelvin, several Mm from the solar surface and is sparse in plasma, but is dominated by magnetic fields, i.e.* $\beta < 1$. *The figure was taken from Lang (2007).*

is constantly monitored using the Solar Dynamics Observatory/Helioseismic and Magnetic Imager (SDO/HMI; see Scherrer et al. 2012) instrument which takes images of the continuum intensity (over the Fe-I line at 6173Å) and the magnetic field via the Zeeman effect (e.g., see Babcock 1967, for details).

The chromosphere sits above the photosphere and stretches to a few thousand km above the solar surface, an example image of the chromosphere taken by the Swedish Solar Telescope (SST; see Scharmer et al. 2003) CHROMospheric Imaging Spectrometer (CHROMIS; e.g., see the discussion by Löfdahl et al. 2021) instrument is presented in fig. 1.5. It is optically thin except in strong spectral lines; strong lines are those with emissions that have a strong intensity and typically have designated Fraunhofer lines, e.g. the C and F lines corresponding

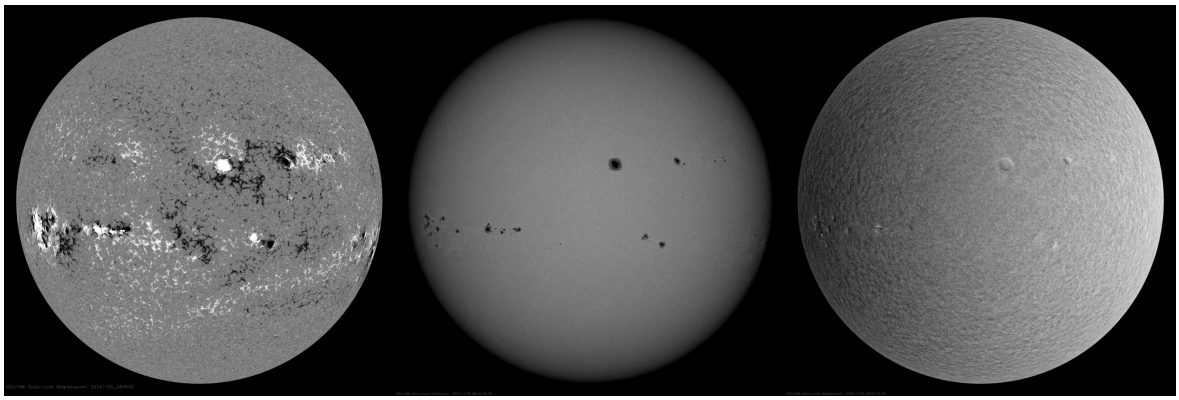

Figure 1.3: *Examples of photospheric images from the helioseismic and magnetic imager (HMI) instrument, onboard the solar dynamics observatory (SDO). Shown on the left are magnetogram images, where dark regions indicate negative polarity of the magnetic field and the lighter regions have a positive polarity. In the centre, the continuum intensity is shown, where the dark regions indicate a drop in intensity. On the right, the Doppler measurements over the disk of the Sun are shown. The lighter regions indicate motion away from away from the viewer; note here that the rotation of the Sun directly influences measurements. Images courtesy of NASA.*

to H-α and -β respectively, and the H and K lines corresponding to different changes in energy levels of singly-ionised calcium, which release photons in the near-ultraviolet at 396.847nm and 393.366nm, respectively (see fig. 1.4). The chromosphere is particularly hard to study due to the spectral lines produced in this region covering a wide range of heights, presented in fig. 1.6.

At the top of the chromosphere lies the transition region in which there is a sharp drop in the density, around 1 order of magnitude, of the plasma and a steep gradient in temperature corresponding to an increase of around 2 orders of magnitude, contributing a large heating to the outer atmosphere, shown in fig. 1.2. This is due to a range of complex dynamics which contribute to the transport of energy and couple it strongly to the next and final layer of the solar atmosphere.

The solar corona extends out from the transition region into the heliosphere, which encloses the solar system. The corona is optically thin in almost all of the electromagnetic spectrum, but is visible in white light during solar eclipses or using a coronagraph. The corona is home to almost entirely transient phenomena, which can be viewed in X-ray spec-

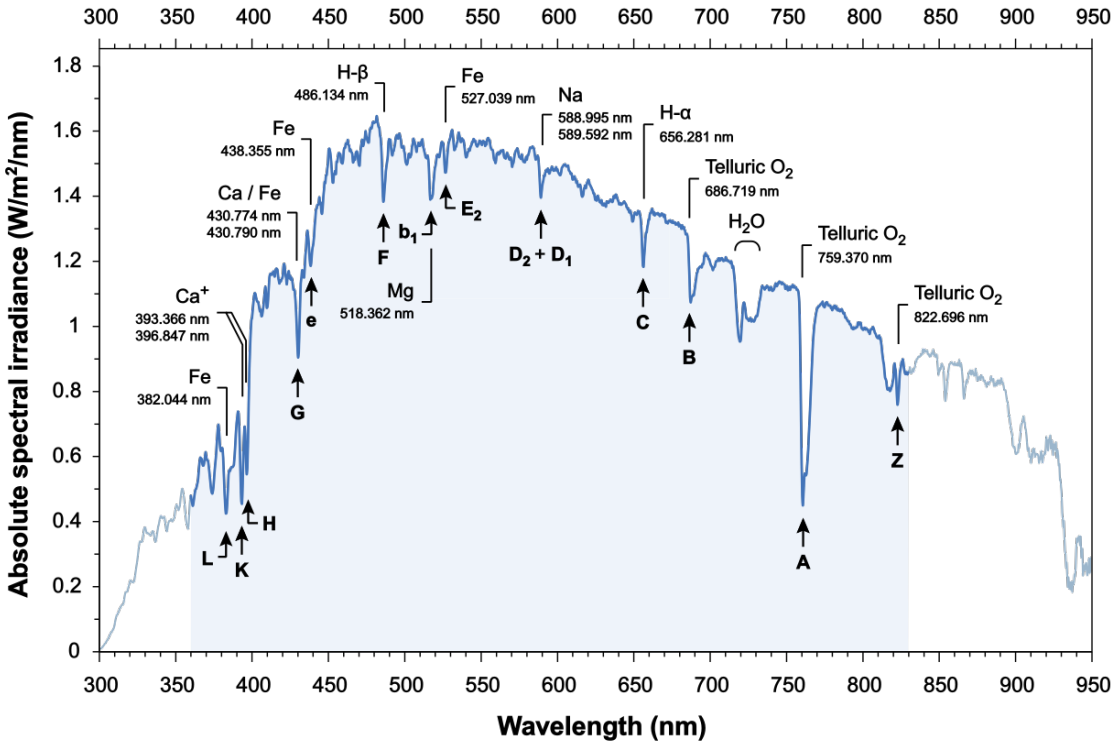

Figure 1.4: *Figure depicting some of the key Fraunhofer lines and the elements responsible for their formation. Each sharp dip in the spectral irradiance highlights a distinct spectral line. Here, the H-α, -β, Fe and Ca lines, are of specific interest due to their correspondence with the solar photosphere and chromosphere. Namely, the H-α and -β lines and the Ca lines are used to observe the chromosphere and the Fe lines are used to study the photosphere. Image taken from Cyamahat (2022)*

.

tra. Observations reveal 3 primary structures in this part of the atmosphere, these are coronal holes, coronal loops and X-ray bright points. The first appears as darkened regions in the solar atmosphere due to a drop in the relative density of plasma, which is transported along open field lines into the solar wind. Coronal loops are large structures consisting of closed magnetic field lines which connect the opposing polarities of ARs in the photosphere. The last of these refers to small loop structures with a typical diameter of 22Mm that appear across the disk of the Sun in great numbers and last only a few hours (Priest 2014).

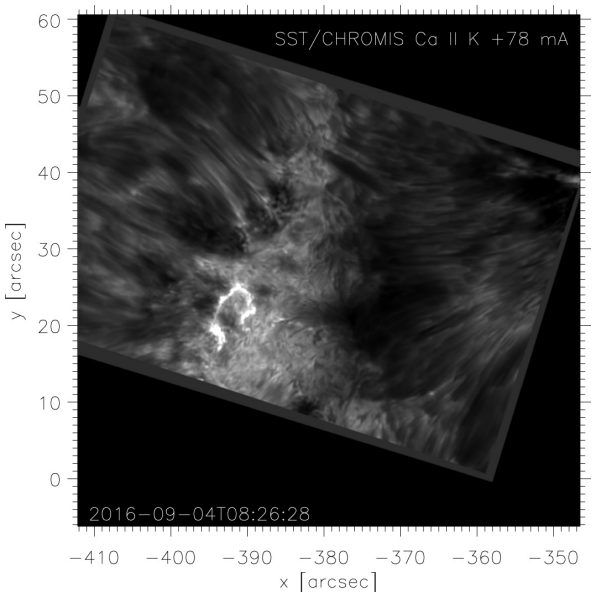

Figure 1.5: *Example of an image of the chromosphere, highlighting a brightening caused by footpoints of a magnetic loop structure, taken in the Ca II (K line of the Fraunhofer spectrum) taken by SST/CHROMIS. Image taken from Testa et al. (2023).*

Active Regions

The structure of the solar atmosphere is highly dynamic, however as described above, we are able to distinguish the atmosphere into layers based on somewhat static properties. This distinction is in fact typical in descriptions of the Sun, which can be considered as the superposition of two paradigms. The first is dubbed the 'quiet Sun' (QS); in this paradigm we consider the Sun to be a large, rotating gaseous body—where properties vary radially out from the core. The second paradigm is more complex and thus significantly more interesting; it describes the 'active Sun' in which the Sun exhibits largely transient behaviour as a result of its ever-changing magnetic field (Priest 2014).

The active Sun covers all transient phenomena such as SFs and CMEs, most of which are driven by the dynamics of the Sun's magnetic field. Magnetic features on the solar surface often manifest as sunspots and pores which can be observed, using appropriate protective equipment, by eye. Any region with a significant level of magnetic flux is termed an active region (AR). ARs are the hubs of notable activity on the Sun and present an opportunity

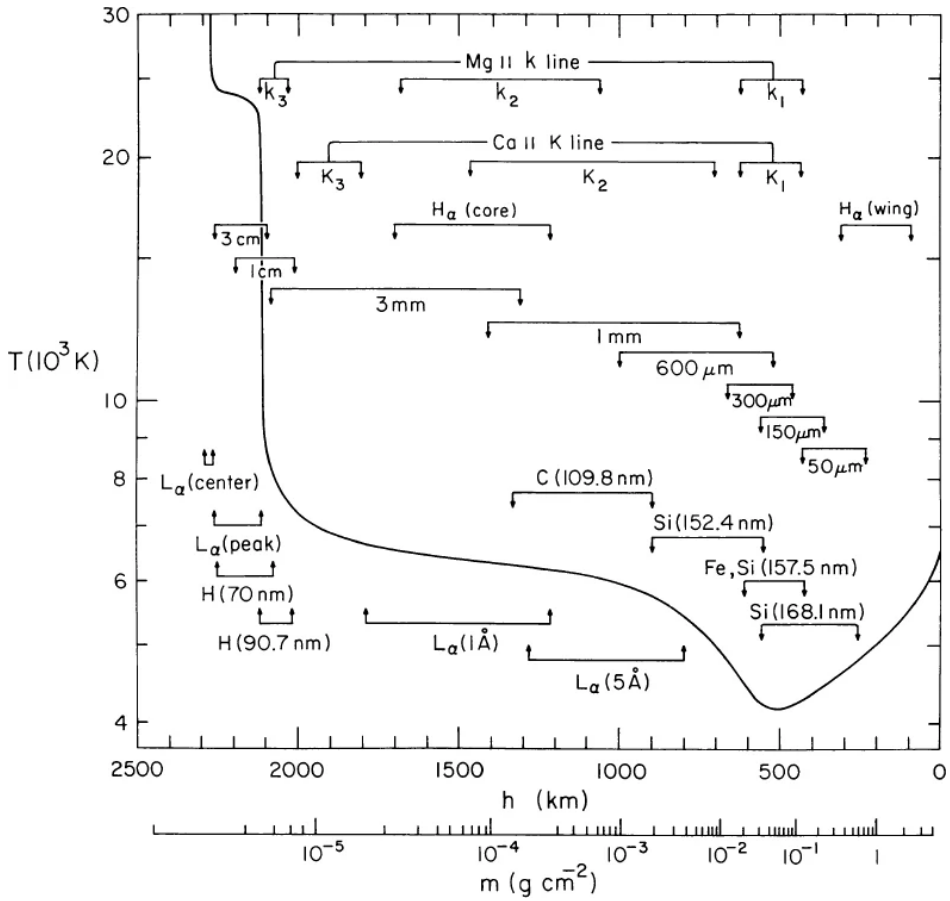

Figure 1.6: *Figure showing the Vernazza, Avrett and Loeser (VAL) model for the solar atmosphere. The axes show the temperature of regions of the solar atmosphere against their respective heights. This highlights the formation heights of spectral lines present in observations. The figure was taken from Vernazza et al. (1981).*

to understand the influence of strong magnetic fields on plasma, which is useful not only in modelling the Sun and SW, but also in the context of energy production, where strong magnetic fields are used to confine plasmas for nuclear fusion.

Active regions, defined by van Driel-Gesztelyi & Green (2015) encompass the totality of observable phenomena represented by magnetic fields from the photosphere up to the corona, visible from radio waves to high-energy X- and γ-rays, which accompany the emergence of strong twisted magnetic flux on the $\geq$ kG scale. This definition is useful but is based on what can be observed during the presence of a sunspot and leaves questions about how the flux emerges and decays, as well as ignoring scale and density. A stronger definition should

discuss preceding physics which appear alongside the AR such as strong up and downflows (several to tens of kms^{-1}) observed through Doppler velocities (see Strous & Zwaan 1999), horizontally diverging flows prior to emergence (see Toriumi et al. 2012) and a minimum lifetime of days (van Driel-Gesztelyi & Green 2015). However, for simplicity and practicality throughout this thesis, the definition of $> 1kG$ magnetic flux will be used when applicable.

ARs form due to strong near-horizontal flux tubes, which are generated by a dynamo action at the bottom of the convective region. These magnetic flux tubes are buoyant compared to the surrounding plasma and thus are convected and twisted through the convective region of the Sun into the photosphere. Once these flux tubes reach the surface, they become more vertical and strong magnetic loop structures form, which often manifest in the visible spectrum as pores or sunspots in the photosphere. These ARs vary largely in their complexity based mostly on the stage of their evolution as well as the strength and twisting of the magnetic field involved in their formation.

Typically, ARs are nested and form on top of pre-existing ARs, thus their evolution is dependent on each other. When magnetic flux emerges at these sites, they carry with them magnetic helicity and free magnetic energy. The magnetic helicity can be thought of as a measure of how entangled the magnetic field lines are, i.e. how complex the topology of the magnetic field is. This puts a lower bound on the energy carried by the magnetic field. As these structures move up through the convective zone to the photosphere, they exert a force on the surrounding plasma, so that as they reach the surface the granules stretch until they fragment and dark spots appear on the surface of the Sun. These dark spots are called pores, or sunspots if they become large enough. These regions are dark due to their relative brightness compared to their surroundings as they contain much less dense plasma, thus the plasma becomes cooler and radiates less heat.

The emergence of magnetic flux, during AR formation, takes several hours to days for the magnetic flux to peak in the atmosphere. Once the magnetic flux has peaked, the decay process begins, which may take several days to several weeks. The decay process is a result of the fragmentation and reconnection of small parts of the magnetic field, which slowly breaks

up the large-scale magnetic field. This decay of the magnetic field happens by a process called magnetic flux cancellation. This is a result of the high magnetic Reynolds number in the photosphere. The magnetic Reynolds number $\mathrm{Re_m}$ describes the ratio of advection to magnetic diffusion in the flow, which depends on the velocity scale and length scale, and the resistivity of the plasma in question. In the $\mathrm{Re_m} \ll 1$ limit, energy is dissipated from the flow via the means of magnetic diffusion. In the opposite $\mathrm{Re_m} \gg 1$ limit, advective processes dominate and thus energy dissipates from the flow via turbulence. Additionally, in the second limit, the magnetic field becomes frozen-in the plasma and thus magnetic structures are determined by plasma motions; this is discussed in further detail in section 1.3.2. Dominance of advection makes dense plasmas easy to study as motions are coherent and correspond to the evolution of the fluid rather than the magnetic field. Equally the frozen-in property of the magnetic field lines is of interest as it makes studying the evolution of the magnetic structure possible through understanding of the evolving flow structures. Thus small regions of magnetic flux are advected and twisted causing small-scale reconnection events in the lower atmosphere, restructuring and cancelling out the magnetic field. Advection is the only process that has a notable influence on the magnetic field as the diffusion time for relatively small sunspots is on the order of ~ 1000 years (van Driel-Gesztelyi & Green 2015).

There are two widely recognised classification schemes used in order to categorise sunspots by a mix of qualitative and quantitative properties. An older, but still frequently used scheme is the one proposed by Hale et al. (1919), which has three basic classes of ARs:

- α: unipolar;
- β: bipolar;
- β-γ: bipolar structure with no continuous dividing line between polarities;
- γ: multipolar.

This scheme was later altered by Künzel (1960) to include the δ variety of sunspots, adding to the scheme the classes

- δ: a multipolar sunspot, in which the different poles share the same umbra.

1.2 Observations of the Sun

Understanding the Sun is difficult, in both observations and theory; a number of reviews present the difficulties in understanding magnetic flux emergence (e.g., see Cheung & Isobe 2014), in understanding magnetic structures in the solar atmosphere (for observations, see Parenti (2014), for theoretical and numerical models see Wiegelmann & Sakurai (2012)) and many others. In recent decades, a number of ground-based and space-based observatories have been created. Ground-based telescopes include the likes of the Swedish Solar Telescope (SST; see Scharmer et al. 2002), Gregor (see Schmidt et al. 2012), the big bear solar observatory (BBSO; see Denker & Wang 1998), the Atacama large millimetre array (ALMA; see Wootten & Thompson 2009) and most recently, the current largest solar telescope, the Daniel K. Inouye solar telescope (DKIST; Rimmele et al. 2020). Space-based telescopes include the solar dynamics observatory (SDO, see Pesnell et al. 2012), solar orbiter (SO, see Müller et al. 2020), the Parker solar probe (see Raouafi et al. 2023), the imaging magnetograph experiment (IMaX; a balloon-borne instrument, which observed the Sun from a high altitude within the Earth's atmosphere. See Martínez Pillet et al. 2011), the interface region imaging spectrograph (IRIS; see De Pontieu et al. 2014) and the Hinode solar optical telescope (see Tsuneta et al. 2008), to name but a few. Observational instruments work by measuring the emission of spectral lines or specific bands of radiation (see Silje 2008, for details).

A number of other useful measurements can be taken using observational instruments. For example, Dopplergrams, images of the Doppler shift across the solar surface, can be produced. Doppler shift measurements can be used to provide the line-of-sight (LOS) component of motions from the Solar surface. These LOS velocities are not exact as shifts in the wavelength of light are affected atomic level effects. Doppler and collisional broadening are caused by random motions and alterations to the excitation state of atoms, respectively, and create an uncertainty in the spectral profile (see the discussion by Bell & Meltzer 1959). These effects are enhanced where density increases in the solar atmosphere, thus increasing the pressure and therefore the number of collisions. When these broadening processes create

asymmetries in the spectral line profile, they can be interpreted as red or blue shifts, causing a mismeasurement of motions (e.g. see Chen, Yajie et al. 2022). Motions across the surface, perpendicular to the LOS, are usually identified by tracking motions of image elements that are a result of the plasma flows. However, tracking image elements to produce optical flows is a problem in itself, which has many approaches, some of which are highlighted in detail in a later subsection. Velocities measured this way cannot ever be truly identified due to the aperture problem, which states that there are multiple solutions for velocities that produce the same apparent dynamics (e.g., see Hildreth 1983). Another alternative to measuring motions is through inversions of the magnetohydrodynamics (MHD) equations, by solving for components of velocity using the full 3D magnetic field vector. The magnetic field vector can be determined through the application of the Zeeman effect using spectropolarimetry, where the polarization of monochromatic light is measured. Recording the proper motion, that is the motion of an object relative to a fixed background, of features on the Sun is not a new concept and has been performed numerous times in the literature (e.g., see Newton & Nunn 1951, Ward 1965). Perhaps the first recorded instance of this are sketches by Galileo of sunspots, which were drawn each day and show the drifting of magnetic features over the disk of the Sun, revealing its rotation (Priest 2014).

1.3 Magnetohydrodynamics in the Sun

The Sun, and that which we observe from it, can be modelled as an electrically conducting fluid. One approach to describing this model mathematically is achieved by imposing Maxwell's equations of electromagnetism on a fluid, which is governed by the Navier-Stokes equations (see, Goedbloed et al. 2004). Such an electrically conducting fluid is the plasma, which has already been used to describe much of the solar atmosphere.

In the following, the derivation of the closed set of MHD equations is presented, which consists of the marriage of electromagnetism and fluids mechanics.

The first set of equations describes Maxwell's laws (see e.g. Maxwell 1865). In their

differential form (presented by Fleisch 2008), these are given as

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{1.2}$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}, \tag{1.3}$$

$$\nabla \cdot \mathbf{E} = \frac{\tau}{\epsilon_0}, \tag{1.4}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{1.5}$$

where $\mathbf{E}$ is the electric field, $\mathbf{B}$ is the magnetic induction (or flux density), which, for simplicity, will be referred to as the magnetic field vector, $\mathbf{j}$ is the electric current density, $c \approx 2.998 \times 10^8 \mathrm{ms}^{-1}$ is the speed of light, τ is the electric charge density and $\epsilon_0 = 1/\mu_0 c^2 \approx 8.854 \times 10^{-12}\mathrm{m}^{-3}\mathrm{kg}^{-1}\mathrm{s}^4\mathrm{A}^2$ is the permittivity of free space in a vacuum, which is related to magnetic permeability of a vacuum $\mu_0 \approx 1.257 \times 10^{-6}\mathrm{Hm}^{-1}$.

Each equation describes a different feature of interacting electric and magnetic fields. Equation (1.2) describes how changing magnetic fields, with respect to time, may generate electric fields. Equation (1.3) describes how electrical currents and changing magnetic fields may generate magnetic fields. The final two equations eq. (1.4) states that sources and sinks of electrical current are produced by regions of electrical charge, and eq. (1.5) is a statement about the geometry of a magnetic field, i.e. there are no magnetic monopoles, like a bar magnet there must always be a north and a south pole to the magnetic field.

We can make a simplifying assumption which still makes the equations valid in most scenarios, that is we consider plasma motions to be non-relativistic, i.e. the characteristic speed of the plasma $v_0 \ll c$. By comparing the magnitudes of the magnetic field and the electric field contributions in eq. (1.3), we can show that the magnitude of the LHS far succeeds the magnitude of the RHS by the non-relativistic condition. Thus the electric field term can be dropped, yielding the result that the magnetic field is a consequence of currents in the plasma. Taking the divergence of this further shows that $\nabla \cdot \mathbf{j} = 0$, i.e. currents in MHD have a net zero divergence, that is they flow in closed circuits.

Maxwell's equations are coupled to the plasma via the magnetic force produced by having

non-static currents which are induced by plasma motions according to Ohm's law which is given by

$$\mathbf{j} = \sigma(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \tag{1.6}$$

where $\mathbf{v}$ is the plasma velocity and σ is the electrical conductivity. Since in the MHD model we care only about the interactions between the plasma and magnetic field, it is convenient to demote the electric field and current density to secondary variables. By using eq. (1.6) we may combine eq. (1.2) and eq. (1.3) to obtain the induction equation,

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \eta \nabla^2 \mathbf{B}, \tag{1.7}$$

where $\eta = 1/\mu_0 \sigma$ is the magnetic diffusivity. This equation links the evolution of the magnetic field to the motions in the plasma and is the heart of MHD. Thus the system no longer depends on the current and electric field which are now given by eq. (1.3) and eq. (1.6), respectively.

The following equations describe the fluid nature of the plasma. Starting with Euler's equation (often referred to as the momentum equation) for fluid motion

$$\rho \left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} \right) = -\nabla p + \mathbf{j} \times \mathbf{B} + \rho \mathbf{g} + \nu \nabla^2 \mathbf{v}, \tag{1.8}$$

where ρ is the plasma density, p is the plasma pressure, $\mathbf{j} \times \mathbf{B}$ is the *Lorentz force* produced by, and acts perpendicular to, the magnetic field, $\mathbf{g}$ is the acceleration due to gravity and ν is the kinematic viscosity. This equation comes from Newton's second law of motion $F = ma$, the left-hand side of eq. (1.8) describes the mass and acceleration whereas the right-hand side describes the sum of all forces acting on a single fluid element.

Note that, as well as the standard forces pressure and gravity driving the fluid, in a magnetised plasma there is an additional contribution from the term $\mathbf{j} \times \mathbf{B}$. This force is non-trivial. It acts perpendicular to the magnetic field on conducting materials and so does not contribute any force along the field line. If we express this only in terms of the magnetic

field according to eq. (1.3), in the non-relativistic limit, we have that the Lorentz force $\mathbf{F}_L$ is written

$$\begin{aligned}\mathbf{F}_L &= \frac{1}{\mu_0}(\nabla \times \mathbf{B}) \times \mathbf{B} \\ &= \frac{1}{\mu_0}(\mathbf{B} \cdot \nabla)(\mathbf{B}) - \frac{\nabla \mathbf{B}^2}{2\mu_0}.\end{aligned} \tag{1.9}$$

Written in this form we see that the Lorentz force has two contributions. The first term on the RHS is magnetic tension force. Like a guitar string after being plucked, this force acts to straighten field line when it has a curvature. The second term is magnetic pressure, which acts across field lines to compress them together/decompress them apart. From this we can define the plasma-β, the ratio of kinetic and magnetic pressures, given by

$$\beta = \frac{2\mu_0 p}{B^2}. \tag{1.10}$$

The plasma-β is important in the context of propagating waves, where perturbations may compress either the fluid or magnetic field.

The continuity equation, which is the mathematical statement that matter is neither created nor destroyed, is key for any continuous fluid and is of great importance in MHD, it is written as follows

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \tag{1.11}$$

Here the terms state that the rate of mass increasing and decreasing is equal to the flow of matter in and out of a given region.

The final equation states how energy in the system flows. In ideal MHD we use the adiabatic gas law $p\rho^{-\gamma} = \text{const}$ where $\gamma = c_p/c_V$ is the ratio of specific heats for heating processes at constant pressure c_p and constant volume c_V. From this, the total derivative w.r.t. time is equal to 0 and expanding this we obtain the equation we desire in the following

form

$$\frac{\partial p}{\partial t} + \mathbf{v} \cdot \nabla p + \gamma p \nabla \cdot \mathbf{v} = 0, \tag{1.12}$$

which does for the pressure/energy of the system what the equation of continuity equation does for the mass. The above set of equations together describes the evolution of an electrically conducting fluid flowing within a magnetic field.

To summarise, the complete set of MHD equations, under the assumption that velocities are non-relativistic, are as follows

$$\rho \frac{d\mathbf{v}}{dt} = -\nabla p + \mathbf{j} \times \mathbf{B} + \mathbf{F}, \tag{1.13}$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \tag{1.14}$$

$$\frac{\partial p}{\partial t} + (\mathbf{v} \cdot \nabla) p = -c_s^2 \rho \nabla \cdot \mathbf{v}, \tag{1.15}$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \eta \nabla^2 \mathbf{B}, \tag{1.16}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{1.17}$$

where the gravitational and viscous forces in eq. (1.8) has been contracted to body forces $\mathbf{F}$ in equation eq. (1.13) to account for other external forces and the sound speed $c_s = \sqrt{\frac{\gamma p}{\rho}}$ has been introduced into eq. (1.12). With these is the equation of state for ideal gases

$$p = \rho \frac{\tilde{R}}{\tilde{\mu}} T, \tag{1.18}$$

where $\tilde{R}$ is the gas constant, $\tilde{\mu}$ is mean atomic weight and T is temperature.

1.3.1 MHD Waves

Understanding the dynamics of the Sun is a question of understanding how energy is transported. There are two major mechanisms for transporting energy that are being studied extensively. These mechanisms are MHD waves (e.g., see Jess et al. 2015) and the other is

advection via MHD flows (e.g., see Tziotziou et al. 2023, Toriumi & Wang 2019).

MHD waves are perturbations that are carried by the plasma and magnetic field. They are a solution to the linearised MHD equations for a small perturbation. The derivation is as follows: assume ideal MHD conditions and include a gravitation force which acts vertically down. Consider a static, homogeneous plasma with boundaries which extend to infinity with a background pressure, density and magnetic field denoted by subscript 0 and small perturbations denoted with subscript 1

$$p = p_0 + p_1, \qquad \rho = \rho_0 + \rho_1, \qquad \mathbf{B} = \mathbf{B}_0 + \mathbf{B}_1, \tag{1.19}$$

respectively. Since the plasma is static we only have perturbations which contribute to the velocity

$$\mathbf{v} = \mathbf{v}_1. \tag{1.20}$$

The result of substituting these into the ideal MHD equations, after some algebra, yields the wave equation

$$\begin{aligned}\frac{\partial^2 \mathbf{v}_1}{\partial t^2} =& c_s^2 \nabla(\nabla\cdot \mathbf{v}_1) - (\gamma - 1) g \mathbf{e}_z (\nabla\cdot \mathbf{v}_1) \\ & - g\nabla v_{1z} + (\nabla \times (\nabla \times (\mathbf{v}_1 \times \mathbf{B}_0))) \times \frac{\mathbf{B}_0}{\mu_0 \rho_0},\end{aligned} \tag{1.21}$$

where c_s is the sound speed of the plasma, for which we seek wave-like solutions of the form

$$\mathbf{v}_1(\mathbf{r}, t) = \mathbf{v}_1 e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}, \tag{1.22}$$

where $\mathbf{k}$ is the wave vector and ω is the wave frequency. This simplifies out calculations as the differential operators become the linear maps

$$\frac{\partial}{\partial t} \to -i\omega \qquad \nabla \to i\mathbf{k}. \tag{1.23}$$

Substituting these into the wave equation yields the dispersion relation with solutions $\omega = \omega(k)$

$$\begin{aligned}\omega^2\mathbf{v}_1 =& c_s^2\mathbf{k}(\mathbf{k}\cdot\mathbf{v}_1) + i(\gamma-1)g\mathbf{e}_z(\mathbf{k}\cdot\mathbf{v}_1) \\ &+ ig\mathbf{k}v_{1z} + (\mathbf{k}\times(\mathbf{k}\times(\mathbf{v}_1\times\mathbf{B}_0)))\times\frac{\mathbf{B}_0}{\mu_0\rho_0}.\end{aligned} \tag{1.24}$$

We recover standard pressure waves in a fluid by ignoring the magnetic field ($\mathbf{B} = 0$) and gravity ($\mathbf{g} = 0$). In this case the dispersion relation eq. (1.24) reduces to

$$\omega^2\mathbf{v}_1 = c_s^2\mathbf{k}(\mathbf{k}\cdot\mathbf{v}_1), \tag{1.25}$$

which implies that the perturbation is in the direction of the propagation, i.e. the wave is longitudinal and has a phase speed

$$v_{\text{ph}} = \pm\frac{\omega}{k} = \pm c_s. \tag{1.26}$$

So the only wave which may propagate is a sound wave, also known as an acoustic wave, that travels in all directions equally.

In order to recover magnetic waves from the dispersion relation, we ignore the effects of fluid pressure and gravity ($p = 0$, $g = 0$). Here the dispersion relation reduces to

$$\omega^2\mathbf{v}_1 = (\mathbf{k}\times(\mathbf{k}\times(\mathbf{v}_1\times\mathbf{B}_0)))\times\frac{\mathbf{B}_0}{\mu_0\rho_0}. \tag{1.27}$$

For practicality, write $\mathbf{B}_0 = B_0\hat{\mathbf{B}}_0$ where $B_0 = |\mathbf{B}_0|$ and define $v_A = B_0/\sqrt{\mu_0\rho_0}$ to be the Alfvén speed. With some algebra the above reveals that $\hat{\mathbf{B}}_0\cdot\mathbf{v}_1 = 0$ and further that

$$(\omega^2 - k^2v_A^2)(\mathbf{k}\cdot\mathbf{v}_1) = 0, \tag{1.28}$$

which has yields two solutions for the dispersion relation.

Firstly, consider the case $\mathbf{k}\cdot\mathbf{v}_1 = 0$, implying that the waves are incompressible and

transverse, then eq. (1.27) yields

$$\frac{\omega}{k} = v_A \cos\theta_B \tag{1.29}$$

where θ_B is the angle between the wave vector and the magnetic field. These are *shear Alfvén* waves. These waves travel along the magnetic field with a maximum phase speed when the wave vector aligns with the magnetic field.

The second solution to eq. (1.28) is $\omega = kv_A$, referred to as the *compressional* Alfvén wave. This solution results in a wave with travels in the $(\mathbf{k}, \mathbf{B}_0)$ plane in a direction perpendicular to $\mathbf{B}_0$ and so has components both along and transverse to the wave vector, this causes changes in both pressure and density. The perturbations of the magnetic field, from the induction equation, are found to be in the $(\mathbf{v}_1, \mathbf{B}_0)$ plane and normal to $\mathbf{k}$ and the Lorentz force is in the direction of $\mathbf{v}_1$. Depending on the angle of the vector with the magnetic field, the compressional Alfvén wave may be purely longitudinal or purely transversal—in the former case, this is equivalent to the shear Alfvén wave.

In real-world scenarios kinetic properties are rarely decoupled from the magnetic properties of a plasma such as in active regions where it is pointed out by Gary (2001), and shown in fig. 1.7, that the plasma-β varies as a function of height and it may very well take a value of plasma-$\beta \approx 1$, i.e. kinetic and magnetic effects are equal.

When plasma-$\beta \gg 1$ then the plasma pressure dominates and effects of magnetic pressure can be assumed negligible. The plasma-β is not constant for fixed regions of the atmosphere but changes locally depending on the surrounding plasma properties. Examples of where this limit is valid is in regions of quiet Sun granulation where convective processes are dominant. In the opposite limit where plasma-$\beta \ll 1$ the magnetic pressure dominates and assume the effects of plasma pressure to be negligible. For example, the plasma-β is low in regions where strong magnetic fields dominates such as active regions in the photosphere, in coronal loops.

In cases where the plasma-β is close to unity we must consider the magnetic field and

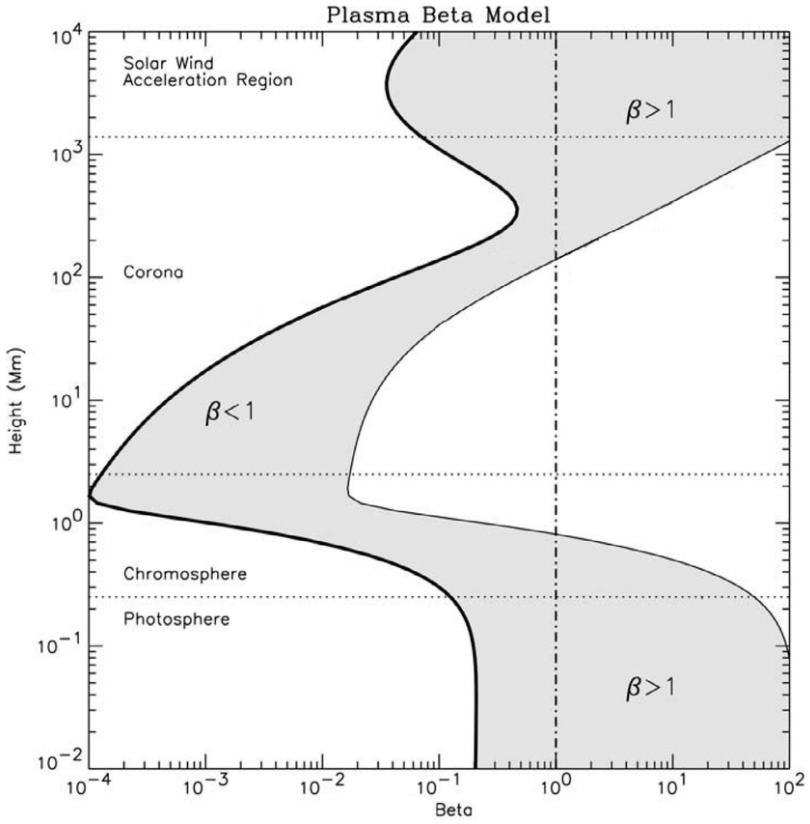

Figure 1.7: *A map of the range of values that the plasma-β may take as a function of height in the solar atmosphere, taken from Gary (2001). The heights shown here are approximate as the location of each region is determined by a number of factors.*

plasma equally. In this case, ignoring gravity, eq. (1.24) becomes

$$\omega^2 \mathbf{v}_1/v_A^2 = k^2 \cos^2\theta_B \mathbf{v}_1 - (\mathbf{k}\cdot\mathbf{B}_1)k\cos\theta_B \hat{\mathbf{B}}_0 + [(1+c_s^2 v_A^2)(\mathbf{k}\cdot\mathbf{v}_1) - k\cos\theta_B(\hat{\mathbf{B}}_0\cdot\mathbf{v}_1)]\mathbf{k}, \tag{1.30}$$

for which the solutions are

$$\omega/k = \left[(c_s^2 v_A^2)/2 \pm \sqrt{c_s^4 v_A^4 - 2c_s^2 v_A^2 \cos 2\theta_B}/2\right]^{1/2}. \tag{1.31}$$

The plus/minus sign in the above shows that there are in fact two phase speeds with one greater than the other. These modes are appropriately named the fast and slow magnetoacoustic waves for the greater and lower phase speeds, respectively.

The magnetoacoustic modes are, like the shear Alfvén wave, dependent on the angle between the wave vector and the magnetic field. Since both the magnetic field and thermodynamical properties are included, we obtain this richer number of modes. The slow wave is a sound wave, which is modified by the magnetic field and the fast wave a compressional Alfvén wave, modified by the gas pressure.

In summary, the solutions of the linearised MHD equations yield three distinct types of wave. One is a purely magnetic wave, which travels through along the magnetic field, called the *Alfvén* wave. This wave travels at the Alfvén speed, which can be considered to be the 'sound' speed of the magnetic field. The other two can be thought of as typical acoustic sound waves broken into two components. The first component is the slow magnetoacoustic wave. Slow magnetoacoustic waves are sound waves which propagate through the plasma, but their propagation is hindered by the magnetic forces. The second component is the fast magnetoacoustic wave, which is a sound wave whose propagation is enhanced by magnetic forces. These waves play a vital role in the heating of the solar atmosphere and restructuring of magnetic fields.

1.3.2 MHD Flows

Whilst the importance of waves in and on the Sun cannot be denied, the transport of energy via plasma flows is deeply complex. Flows largely contribute to the heating of the solar atmosphere via the radiative diffusion of heat from hot plasma convected through the solar interior. Plasma flows are also responsible to the restructuring of the topology of magnetic fields via their advection.

The behaviour of MHD fluids can be broadly characterised by identifying the dominant mechanism for material transport. In fluid mechanics, the Reynolds number compares the ratio of inertial forces to viscous forces; similarly, we can define the magnetic Reynolds $\mathrm{Re_m}$, which compares the ratio of advection to diffusion in the induction equation. Therefore the

magnetic Reynolds number $\mathrm{Re_m}$ is defined as

$$\mathrm{Re_m} = \frac{|\nabla \times (\mathbf{v} \times \mathbf{B})|}{|\eta \nabla^2 \mathbf{B}|} \approx \frac{uL}{\eta}, \tag{1.32}$$

where u and L are the characteristic speed and length scale of the plasma. In an ideal plasma the fluid behaves as a perfect conductor, hence $\eta = 0$ meaning that advection of magnetic flux by the plasma is the only mechanism contributing to the evolution of the magnetic field (Alfvén 1942). In a non-conducting plasma the resistivity $\eta \gg 1$ means that the magnetic field diffuses through the fluid and is not strongly influenced by the flow. Most plasmas on the Sun are highly conducting and therefore diffusion is only expected to be relevant on short scales, so that the advective term becomes negligible. Thus this limit is of limiting interest.

The transport of magnetic energy by plasma motions is a direct consequence of the induction equation. In particular, in the ideal plasma limit $\mathrm{Re_M} \to \infty$ we may ignore the effects of magnetic diffusion. Thus the induction equation simplifies to

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}). \tag{1.33}$$

This form of the induction equation takes on the role of a continuity equation for the magnetic field. Rewriting the double cross product, as dot products, yields

$$\frac{\partial \mathbf{B}}{\partial t} = \mathbf{v}(\nabla \cdot \mathbf{B}) + (\mathbf{B} \cdot \nabla)\mathbf{v} - \mathbf{B}(\nabla \cdot \mathbf{v}) - (\mathbf{v} \cdot \nabla)\mathbf{B}. \tag{1.34}$$

By applying Gauss' law $\nabla \cdot \mathbf{B} = 0$ the remaining terms describe the evolution of the magnetic field as the combination of three effects:

- $\mathbf{B}(\nabla \cdot \mathbf{v})$: the magnetic field being altered by plasma flowing into and out of a volume;
- $(\mathbf{v} \cdot \nabla)\mathbf{B}$: the magnetic field being advected along lines of plasma flow;
- $(\mathbf{B} \cdot \nabla)\mathbf{v}$: the plasma being advected along magnetic field lines.

Hence, from this form of the induction equation we see that the flow is advected along field

lines and as well as being influenced by the Lorentz force when carrying current perpendicular to the magnetic field.

Unlike the mass which is being moved by flow, the magnetic field is not a conserved quantity. However the magnetic flux, which can be loosely defined as the number of magnetic field lines passing through a closed surface, is conserved. That is, the magnetic flux within a volume is deformed equally with the motion of plasma. This result is known as Alfvén's theorem or the 'frozen-in flux' theorem due to the magnetic field lines being frozen into a volume.

Within the Sun, magnetic fields are generated and compressed together to form magnetic flux tubes, as a result of the dynamo action produced by magnetoconvection. Flux tubes are regions that contain dense bundles of magnetic field lines that are tangent to the surface of the volume they enclose. A number of works have already established that the magnetic field distribution across the photosphere exhibits a strong correlation with coherent plasma flow structures (e.g., see Chian et al. 2019, Silva et al. 2023).

1.3.3 Modelling the Sun

Supplementing the great number of observing instruments and therefore wealth of data being collected, there has been a rapid improvement in computational power. This has brought about the development of a number of numerical codes for simulating different aspects of the Sun in great detail. Simulations provide a numerical laboratory for understanding the Sun. Solving the equations that govern the physics of the Sun allows for the study of a full set of information pertaining to the Sun at every single grid point, allowing for a greater insight. For example, fully 3D velocity vectors can be determined throughout an entire simulation box directly. This is important simulations can provide realistic models including synthesised images which closely match real-world observations of the Sun.

A number of numerical codes have been produced to study different regions and effects within the Sun and its atmosphere. These simulations range from modelling the convection of plasma below the surface that forms solar granules in the photosphere, to the modelling

of ARs and complex magnetic field structures in the corona and beyond. To name but a few notable simulations over the last few decades there are the STAGGER (Stein & Nordlund 2012, Stein et al. 2024), MURaM (Vögler, A. et al. 2005, Rempel & Cheung 2014), MANCHA (Khomenko & Collados 2006), Bifrost (Gudiksen et al. 2011), and most recently the R2D2 (Hotta et al. 2019, Hotta & Iijima 2020) which was developed from the MURaM code, and RAMENs (Iijima et al. 2023) which can model patches of the Sun from the convective region out to the solar wind. More detailed descriptions of key simulations are given in chapter 2 and appendix B.

1.4 Flow Recovery methods

The current lack of ability to make any direct measurement of plasma motions in the solar atmosphere limits the analysis that we can perform in on the solar atmosphere. Without efficient and robust methodologies the a large portion of data remains inaccessible and there is a need to produce such methods. To meet this demand, a number of algorithms have been developed. These algorithms can be split into distinct categories based on their approach to calculating the velocity field, these are optical tracking, physical inversions and machine learning (ML).

1.4.1 Traditional Flow Recovery

Optical flows are flows which, when applied to an initial image, advect the pixels in such a way that it transforms it into a second image. Optical methods are simple to understand and can be deployed in a number of ways, thus they are extremely popular and used in a number of scenarios, for example tracking solar granulation (Löptien, B. et al. 2016) as well as AR inflows (Löptien et al. 2017). Algorithms that identify optical flows, by performing calculations to determine a velocity field for transforming images into one-another, present as possibly the most popular method for estimating flows across the solar disk. The algorithm for producing an optical flow is essentially 4 steps which consists of

- Identify a feature in the first image;

- Identify the same feature in the second image;
- Calculate the displacement between the features in the two images;
- Using the spatial and temporal displacement of these features, calculate the velocity vector.

A number of popular optical flow tracking methods exist, namely local correlation tracking (LCT: November & Simon 1988), Fourier-based LCT (FLCT: Fisher & Welsch 2008), feature/coherent structure tracking (FT and CST: Strous 1995, Roudier et al. 1999, respectively) and balltracking (Potts, H. E. et al. 2004). All of the prior listed methods track objects, either pixels or more well-defined structures, between images over a time window to produce a velocity map by measuring the distance travelled by the object. It was shown by Rieutord et al. (2001) that these optical flows correlate well with the proper motions of mesogranular (> 2.5Mm) structures averaged over ≈ 0.5hours. The exact way that a feature may be identified differs from algorithm to algorithm.

The local correlation tracking (LCT; see November & Simon 1988) method works by correlating features within subimages between two time-consecutive frames. In order to calculate the cross-correlations, filtered intensity images are displaced by a vector, relative to each other, and then their product is taken with an apodising window. The integral of this product taken over the size of the window gives the cross-correlation at a central position. Maximising this correlation provides the best description of the displacement of the pixels in the image and therefore provide the best velocity field. Variations of this algorithm exist, namely Fourier-based LCT (FLCT; see Fisher & Welsch 2008) and Induction LCT (ILCT; see Welsch et al. 2004). The former, FLCT, is almost identical to the LCT in methodology, but performs calculations using the fast Fourier transform (FFT) to improve the calculation time of the algorithm and utilises a Gaussian for the apodising window. The choice of the apodising window is due to the FFT nature of FLCT, which generates high frequencies at edges of images where gradients are enhanced. The Gaussian smooths gradients and thus reduces noise generated by the method. The ILCT method uses magnetogram data to

produce velocities whose normal component is consistent with the MHD induction equation (see eq. 1.7) in the ideal limit. The overarching principle of correlation tracking is that the maximised correlations of displaced pixels between two images should describe a fluid motion which advects the constituents of one image into the other.

Similar to correlation tracking, there are other optical tracking methods, which directly track the features in the images. One is aptly named feature tracking (FT; see Strous 1995) and the other is coherent structure tracking (CST; see Roudier et al. 1999), an improved version of the FT algorithm. FT and CST segments an image into objects, which similar to the underlying principle of correlation tracking, are expected to be advected by a flow field into the objects present in the next image. By identifying features between images a flow field can be derived. For example, solar granules in the photosphere are expected to behave like tracers, advected by flows across the photosphere. By identifying granules between images, one can trace a path and therefore a velocity field.

In addition to this, there are other optical methods which identify flows by mapping images to a surface whose height is proportional to the brightness of the image at each point. One such method is the balltracking method presented by Potts, H. E. et al. (2004). In principle, this method works by allowing tracers (balls) to be advected by the corrugations in a surface which apply a force to them. The tracers in the flow will fall into the minima of the surface and pushed around by the brightest objects, for example granules, and therefore the average motions of the tracers will be representative of true flow field. The resultant velocity field from this method is consistent with the LCT methods (Potts, H. E. et al. 2004).

It should be noted that the authors of the LCT algorithm presented that there are a number of issues with the LCT approach (see Simon et al. 1995), namely that even when the signal-to-noise ratio is high through means of spatial and temporal averaging of velocity fields, the results are discouraging, providing flows that are $\approx 1/3$ of what they should be, which was later highlighted again by Verma et al. (2013). The article from Simon et al. (1995) states that results from the LCT methodology needs to be studied carefully and validated properly before use.

A distinct problem with these optical approaches, aside from balltracking, is that they demand a distinct change in the position of features in the images to determine a non-zero velocity. All of the methods listed above perform the best on highly structured images. The requirement of the clear advection of features makes tracking features on small temporal and spatial scales prone to inaccuracy due to apparent gaps in the predicted velocity fields, where no motion is detected. In order to subvert this limitation, the individual velocity fields produced by the methods are averaged temporally and often smoothed spatially to produce complete and continuous flow maps, which produce good correlation with the proper motions at the larger scales, described above.

As mentioned previously, and alternative to using optical tracking to identify flows, one can use physical inversions to produce a velocity field. For example, the differential affine velocity estimator for vector magnetograms (DAVE4VM; see Schuck 2008) uses magnetograms from observation in order to produce a velocity field that is consistent with the ideal MHD induction equation.

The final example of methodologies is machine learning. For example, the DeepVel (DV; see Asensio Ramos, A. et al. 2017) artificial neural network (NN) which can be trained on images synthesised from simulations, where the velocity field is known, in order to construct a model that will output velocities from new images. The details of this network will be described in the following section.

1.4.2 Neural Networks

In the wake of high performance computing and big data, the task of categorising, understanding and inferencing data has fallen onto computers. The process of computers constructing models for inferencing from data is called machine learning. One category of machine learning is supervised learning, where an algorithm is established to generate a mapping between input data and a labelled output. NNs excel at supervised learning tasks. They attempt to mimic biological NNs, i.e. the brain, in order to solve supervised learning tasks. They achieve this by forming layers of neurons, which form non-linear combinations of the inputs.

Developing NNs has been of interest for the best part of the last century (McCulloch & Pitts 1943). Rosenblatt (1958) described the perceptron, a basic NN with a single unit that weights inputs to produce a binary output to solve linear problems, and a learning algorithm was presented on how to adjust the parameters of the perceptron. This output, which is based on the weighted sum reaching a threshold, is called an activation function and can take many forms. Perhaps one of the most common examples is the sigmoid activation function (Rumelhart et al. 1986, McCulloch & Pitts 1943), and in modern times much success has been found with the rectified linear unit (ReLU) activation function (Nair & Hinton 2010).

Networks with single layers are limited since they could only be used to solve linearly separable problems (Minsky et al. 2017). However, it was shown that by layering neurons, NNs could be used to approximate any non-linear function (Hornik 1991), this, in tandem with the back propagation algorithm (Rumelhart et al. 1986) allowed for the development of large scale networks which could solve more complex problems. The structure of NNs can be broadly split into 3 categories: the input layer, hidden layers and the output layer. The input and output layers behave as their name suggests with the input layer taking in some data and the output layer spitting out the expected label associated to the data. The hidden layers of neurons sit between the input and output layers, their task is to weight the connections of neurons to build a successful mapping from the first and final layers. Any network which contains more than one hidden layer is often referred to as a deep NN.

Typical examples of hidden layers are activation layers whose task, like the perceptron, is to produce an activation signal when a weighted sum reaches a threshold. Another type is a convolution, in which values for the specific form of the convolution are optimised to identify features that are deemed important by the network. For example, it is often the case that the individual inputs to a network, say the pixel values in an image, are not particularly useful, but groups of inputs will provide useful context. To this end, convolutional NNs (CNNs; see Lecun et al. 1998) excel at tasks where recognising patterns in data is important, such as identifying granules or solar pores. The convolutions group together units of data, such as pixels in an image, in order to reduce the number of parameters. An advantage of CNNs is

that they have translational invariance, that is once a pattern in the data is learned, they can then recognise the same patterns regardless of where it appears. However, it should be noted that this only up to translation—these features should share the similar spatial scales and magnitudes as those identified in the training data.

NNs typically contain many parameters, all of which needs to be optimised to perform a task. Networks optimise these parameters by using a weighting system called the loss function. The loss function (sometimes referred to as the cost function) is a proxy for the success of the predictions; where the smaller the value of this function, the better the predictive performance of the network. The loss function can be used to determine how parameters of the network should be altered through means of gradient descent (GD; e.g., see Rosenblatt 1958). The GD algorithm is computationally expensive as it requires working out gradients for the weighting of every connection in the network. The back propagation algorithm was thus developed and was a key step forward in the evolution of ML as it calculates every gradient as a series of linear operations of the gradients in the last layer.

1.4.3 DeepVel

Recently, DeepVel (see Asensio Ramos, A. et al. 2017), a fully convolutional deep neural network, has been developed for the use of recovering two- or three-dimensional pixel-to-pixel velocity fields little using as little as two solar intensitygrams, magnetograms etc. It presents advantages over other existing methods as it can be trained to give excellent performance on specific datasets, where inputs require no spatial or temporal smoothing and averaging like commonly employed correlation tracking techniques. DV learns to inference velocities using data from numerical simulations of the Sun. As well as efficiency concerns, this method presents an additional advantage over others by being able to rapidly emulate the physics of flows that have been calculated in realistic numerical models at the cost of a long initialisation period for training the network.

The architecture of DV includes several types of layers, including convolutions, ReLU activations, batch normalisations and sums. The ReLU layers consist of activation functions

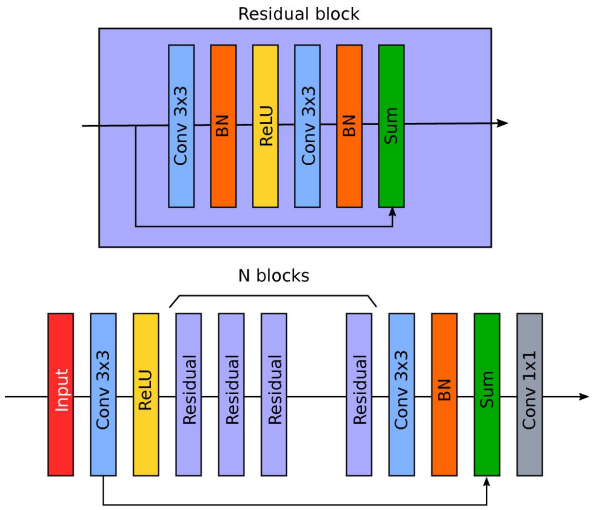

Figure 1.8: *Sketch of the architecture of the DeepVel neural network from Asensio Ramos, A. et al. (2017) showing the layers and transformations acting on an input (left) to the output (right). The highlighted block reveals the combination of layers that form the residual blocks of the network.*

that take a value of 0 for negative activations or take a value equal to the activation otherwise, thus effectively switching off the neuron rather than giving it a negative contribution which helps encourage other neurons to learn and lessens the influence of potentially unhelpful pathways. Another advantage to these functions is that they have a constant gradient in the positive region, preventing them from exploding or vanishing and hence improving learning performance (e.g., see Agarap 2019). Batch normalisation changes inputs to the next layer to have mean 0 and unit variance, which increases the speed of convergence (e.g., see Ioffe & Szegedy 2015). These layers are connected together into residual blocks, which therefore compress and simplify data to make more meaningful and useful connections, and speed up the training process. A complete diagram of the network architecture is provided in fig. 1.8 and full details of the network architecture and design choices are given in Asensio Ramos, A. et al. (2017).

DV has shown success in recovering velocity fields from as little as two consecutive images (Tremblay et al. 2018). The network performance has been shown to be at least as good as FLCT (Asensio Ramos, A. et al. 2017, Tremblay et al. 2018) in reproducing spatially and temporally averaged velocity fields from simulations of QS convection. In addition to this,

some success has been shown for the application of DV within ARs (e.g., see Tremblay & Attie 2020).

1.5 Lagrangian Flow Analysis

Lagrangian analysis is important as objectivity should be dealt with rigorously when working with observations of a fluid flow on a distant rotating mass. Without careful analysis, even in the most simple settings, flows and dynamic behaviour can be erroneously identified with ease. Perhaps the simplest example, highlighted by Tziotziou et al. (2023), is in the case of the Couette flow where fluid flow runs parallel to the axis of a pipe and has a magnitude proportional to the perpendicular distance from one wall of the pipe. This type of flow generates a shearing effect. Suppose the pipe has an axis along the y-direction and has infinite depth in the z-direction. In this setup, the flow $\mathbf{v}$ is described by

$$\mathbf{v} = (0, \xi x, 0). \tag{1.35}$$

Taking the curl of the velocity field, to measure the vorticity ω, yields a non-zero (assuming $\xi \neq 0$) component along the z-axis,

$$\boldsymbol{\omega} = (0, 0, \xi). \tag{1.36}$$

This non-zero vorticity may be misidentified as a rotational flow; a rotation would be observed by placing a rod, perpendicular to the pipe's axis, within the flow. However, it is easy to see that the resulting flow does not generate any swirling motion; injecting a dye into the flow, instead of a rod, would instead reveal streaklines that are straight. Thus, in order to properly record the dynamics of flow processes, a more objective approach is needed.

The Lagrangian approach to motion provides a more rigid framework to characterise fluid motions as it is concerned with following the frame of the fluid elements as opposed to the motion of fluid elements within a frame. As in the example of the Couette flow above, a fluid element moving with the flow will simply become elongated, characterising the shearing motion that is present in the flow. A number of these examples are presented by Haller

(2015*a*).

1.5.1 Coherent Structures and Lagrangian Coherent Structures

In a flow, a coherent structure is a discernible pattern in a flow which persists on a sufficiently long temporal or spatial scale. Examples of such structures include saddle points and vortex motions within a flow, all of which are material surfaces. Identification of coherent structures enables us to characterize the local behaviour of flow. Material surfaces that are identified by tracking the flow are called Lagrangian coherent structures (LCSs) and they partition regions that exhibit similar dynamics, orchestrating the flow into discernible coherent patterns. The transport barriers are determined by advecting particles along a time series of the flow. Thus, the LCS theory facilitates a novel and precise analysis of flow properties by identifying surfaces that locally maximize attraction, repulsion, and shearing (Haller & Yuan 2000, Haller 2015*b*). In other words, LCSs provides a skeleton for the plasma flow dynamics, decomposing complex flow behaviours into dynamic building blocks.

1.5.2 Finite-Time Lyupanov Exponent

The Finite-Time Lyupanov Exponent (FTLE) is a value which can be used in order to define LCSs by seeking repelling and attracting barriers in a flow field. By advecting particles in a flow, one can measure the distance between particles in the flow as time progresses. The final distance after some time is used to calculate the FTLE, which then determines the amount of stretching that has taken place between two particles inside the flow. Since FTLE tracks particles that move with the flow, the barriers that are formed also move with the flow, thus providing a Lagrangian structure.

These structures describe regions where the flow is locally repelling and attracting, thus they behave as barriers over which fluid elements cannot cross. Fluid elements either move toward or away from these barriers without ever passing through them. As such, the barriers provide a skeleton for how the flow field as a whole behaves. In order to identify the attracting and repelling LCSs, a uniform grid of particles is advected over a time-dependant velocity

field. Consider $\mathbf{v}$ as a velocity field in a spatial domain and $\mathbf{x}(t_0) = \mathbf{x}_0$ as a fluid particle at the initial position that follows a trajectory in space given by solutions of the initial value problem. Then,

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}(\mathbf{x}, t). \tag{1.37}$$

The displacement of particles at time $t = t_1$ from its initial position at time $t = t_0$ can be found as:

$$\mathbf{x}_1 - \mathbf{x}_0 = \int_{t_0}^{t_1} \mathbf{v}(\mathbf{x}, t)dt. \tag{1.38}$$

Then, the flow map can be defined as:

$$\mathbf{x}(t_1) = \phi_{t_0}^{t_1}(\mathbf{x}_0), \tag{1.39}$$

where the operator $\phi_{t_0}^{t_1}(\mathbf{x}_0)$ maps the initial position of a particle at $\mathbf{x}(t_0)$ to a final position $\mathbf{x}(t_1)$. In the 2D case, particles are given initial positions $\mathbf{x}_{i,j}(t_0)$, where i and j are indexes of the particle's initial position relative to a chosen centre point. Distances between the central particle and its 4 nearest initial neighbours, on a 2D 5-point stencil, are tracked as they are advected by the flow. The stencil describes a central particle surrounded by 4 particles on the left, right, top and bottom of the centre on a grid.

The deformation gradient after the advection is given by the Jacobian matrix of partial derivatives of the flow map

$$D\phi_{t_0}^{t_1}(\mathbf{x}_{i,j}) = \begin{pmatrix} \frac{x_{i+1,j}(t_1) - x_{i-1,j}(t_1)}{x_{i+1,j}(t_0) - x_{i-1,j}(t_0)} & \frac{x_{i,j+1}(t_1) - x_{i,j-1}(t_1)}{y_{i,j+1}(t_0) - y_{i,j-1}(t_0)} \\ \frac{y_{i+1,j}(t_1) - y_{i-1,j}(t_1)}{x_{i+1,j}(t_0) - x_{i-1,j}(t_0)} & \frac{y_{i,j+1}(t_1) - y_{i,j-1}(t_1)}{y_{i,j+1}(t_0) - y_{i,j-1}(t_0)} \end{pmatrix}. \tag{1.40}$$

This matrix is used to compute the Cauchy-Green deformation tensor,

$$\Delta = \left[D\phi_{t_0}^{t_1}(\mathbf{x}_{i,j})\right]^T D\phi_{t_0}^{t_1}(\mathbf{x}_{i,j}),$$

where the superscript T denotes the matrix transpose. Finally, the FLTE field is calculated

by means of

$$\mathrm{FTLE}_i^{t_1}(\mathbf{x}) = \frac{1}{|\delta t|} \ln \sqrt{\max(\lambda_i)}, \quad i = 1, 2 \tag{1.41}$$

where λ_i are the eigenvalues of Δ.

Given a time interval of length δt, particles are integrated forward in time over the interval $[0, \delta t]$ and integrated backwards in time over the interval $[-\delta t, 0]$, to produce the forward-FTLE and backward-FTLE fields. Ridges formed by the largest FTLEs describe the most strongly repelling structures in the plasma flow in forward time, the most strongly attracting structures in backward time.

A number of important findings related to the FTLE and solar observations are presented in later chapters. Whilst this method is deployed numerous times, the focus of this thesis is not on advancing the theory of these methods, but the novel applications of existing methods and validation of them using these advanced Lagrangian and other similar techniques.

1.6 Thesis outline

This thesis evaluates the suitability of the convolutional neural network DeepVel for recovering plasma flows in magnetically dominated regions of the solar atmosphere. Thereby, answering the question of how plasma flows in the solar atmosphere can be recovered accurately and used to understand the dynamics in evolving regions of concentrated magnetic flux.

The contents of this thesis aims to address a number of points in assessing flow recovery performance and applicability of DeepVel to studying the solar atmosphere in some of the most complex scenarios. The thesis is therefore organised as follows:

- Chapter 2 addresses how flow in the solar atmosphere can be identified by comparing DeepVel to the established FLCT method. DeepVel was trained on a high-resolution simulated data capturing the evolution of a flux tube under magneto-convection, allowing it to learn complex flows in the photosphere. DeepVel has never previously been used to identify flows in regions containing realistic strong flux emergence. Additionally, no previous studies have aimed to identify how DVs performance on new data

(e.g. observations) will differ from data used for training. Whilst a proper ground truth set for flow data on the Sun is missing, a second simulation with a different initial setup provides the ideal testing ground for checking how well DV generalises to new data. Training and cross-validating with simulations with similar properties provides a challenge in ensuring that the recovered flows are not affected by overfitting to specific boundary conditions. Various metrics assess its physical accuracy and advantages over FLCT.

- Chapter 3 tests DeepVel's ability to identify Lagrangian coherent structures in photospheric flows, addressing the ability of neural networks to emulate the physics of simulations and ability to generalise to new scenarios, as well as highlighting a novel method for studying flows within ARs. These structures define the skeleton of plasma motion, serving as a novel method for evaluating DeepVel's predictions. By applying the model to a second simulation, the study examines error propagation and its potential for analysing flow dynamics in observational data. Both chapters 2 and 3 are summarised in the work of Lennard et al. (2025).

- Chapter 4 introduces a novel method for detecting strong magnetic flux concentrations using horizontal flow fields—thus highlighting how new and accurate flow information, derived by neural networks, can be used in the way of forecasting and tracking emerging magnetic fields on the Sun. Finite-time Lyapunov exponents are computed across an active region simulation, revealing correlations between flow structures and high concentrations of magnetic flux. Thus, highlighting a novel classification method for identifying active regions and assessing DeepVel's role in magnetic structure detection, with implications for space weather forecasting.

- Chapter 5 indicates the ability of DeepVel to learn flows from another region of the solar atmosphere. DeepVel's initial design was intended to be used on the solar photosphere where flow structures and images are largely coherent thanks to the density of the plasma, which contributes to it being optically thick and hence easy to track features

within. In order to progress past this, DV was used to identify coherent swirling motions in the solar chromosphere, where the plasma is more sparse and therefore contains many incoherent flows, making learning of flows tricky. DeepVel was tested by determining the presence and location of swirls by employing a vortex detection method. The results show a proof of concept extension of DV's application to the solar chromosphere where it is able to identify coherent structures like swirls.

- The final discussion addresses key findings and future refinements including the next steps in moving toward applying these methods successfully on observations.

Chapter 2

Comparing DeepVel and Fourier-Local Correlation Tracking for Flow Recovery in Simulated Active Regions

2.1 Introduction

Analysing high-resolution solar atmospheric observations requires robust techniques to recover plasma flow features across different scales, especially in active regions. Current methodologies often fall short in capturing subgranular-scale flows, and there is limited research on the errors introduced by the recovered velocities in analysing the properties flows in the presence of intense magnetic flux. This chapter concentrates on validating the effectiveness of DV in recovering subgranular and mesogranular-scale within and surrounding active regions versus another well-established method as a benchmark.

Presently, observations are restricted to line-of-sight (LOS) measurements of plasma velocity via the Doppler effect. Recovery of the horizontal velocity fields from observational data is a challenging task (see Rempel et al. 2022) and several techniques have been proposed

to deal with this problem (Tziotziou et al. 2023). Flow recovery techniques can be broadly split into three distinct categories, these are optical flow tracking, physical inversions and neural networks and are highlighted in chapter 1.

The study by Tremblay et al. (2018) compared the effectiveness of DV with the optical flow methods LCT, Fourier-based LCT (FLCT) and coherent structure tracking (CST). These methods were applied to synthetic photospheric observations at a resolution of the SDO/HMI instrument ($\delta x = 0.5$"pix^{-1}) averaged over 30 minutes. DV was able to reproduce velocities best at granular and subgranular scales in the quiet Sun (QS). However, FLCT performed better at the mesogranular scales. Due to the averaging of velocities, the capabilities of the methods are limited to seeing features that are $\approx 1/2$ the size of granules and transient features which last below 30 minutes are lost. With this and the fact that DV is limited to QS data, we are unable to see the full capabilities of DV and the range of scenarios it may be applied. Later studies improved the previous ones by introducing a new architecture to DV (Tremblay & Attie 2020) and testing this and the original DV architecture on recovering flows from a sunspot at the same resolution as HMI (Tremblay, Benoit et al. 2021). In both instances, only the 30-minute time-averaged velocity fields are considered in the analysis, despite DV being trained on significantly higher cadences of $\approx$ 45s. The sunspot used in the AR study is from a MURaM simulation (see Rempel 2015), which presents the decay of an axis-symmetric cylinder of magnetic flux embedded vertically in the box that covers 18Mm depth. After an initial relaxation period of around 5.5h the flux tube is left to decay under the influence of magnetoconvection, however due to the nature of the setup, only a partial decay takes place over the 100h runtime of the simulation. With this and the fact that the DV network was only tried and tested on the same simulation used for training, there is still a gap in understanding the applicability of DV to new data at finer resolutions (< 1") and in environments, which present a realistic evolution/decay of intense magnetic flux in the solar atmosphere.

In addition to the simulated sunspot presented by Rempel (2015), using the MURaM code, there exists a wealth of simulated solar data from myriad codes. Examples of these

numerical simulations are the MURaM (Vögler, A. et al. 2005), CO5BOLD (Freytag et al. 2012), Stagger (Stein & Nordlund 1998), Bifrost (Gudiksen et al. 2011) and R2D2 (Hotta & Iijima 2020) codes. These codes present working models of magnetoconvection up to the photosphere and beyond. Some of these codes have been modified to simulate the evolution of a strong magnetic flux of varying initial configurations. The MURaM simulation by Rempel & Cheung (2014) introduces an untwisted semi-torus of magnetic flux into the convective region by advecting it through a bottom boundary, $<$ 16Mm below the photosphere. The paper by Bjørgen, Johan P. et al. (2018) highlights simulations of an AR using the Bifrost code, which imposes an pre-formed bipolar AR into the uppermost layers of the convective region. These simulations, and others in the literature, have opted to model magnetoconvective processes using a shallow box that only covers up to $\approx$ 16Mm below the surface (see, e.g. Chen et al. 2023, Beeck, B. et al. 2012, Rempel et al. 2009, Rempel & Cheung 2014). These setups give only a short time for the magnetic field to evolve naturally under the influence of magnetoconvection leaving limited information on the effects of the interplay between plasma and strong magnetic fluxes.

A more recent model for flux emergence was presented by Hotta & Iijima (2020), in which the entire convection zone (200Mm depth) is simulated. This allows an initial force-free magnetic flux tube, positioned at a depth of around 35 Mm below the surface, to evolve under a more realistic boundary condition than simulations with shallow boxes. The simulation starts with a twisted magnetic flux tube, where the force-free parameter $\alpha = 2.43 \times 10^{-7} \mathrm{m}^{-1}$, with $>$ 10kG strength is placed at a depth of $\approx$ 35Mm depth below the photosphere. In this simulation the α is determined as a ratio of the first root of the Bessel function J_0 and the characteristic length scale of the flux tube, chosen to be the radius. This simulation excludes the effects of physical resistivity, i.e. assumes the plasma to be physically ideal, however, the minimum scale in our case is the pixel size, 96km. Khomenko et al. (2014) highlights that the effects of physical diffusivity and resistivity are negligible at the scales we are interested in, only having a significant influence at scales around 100m in both quiet and active Sun conditions at photospheric heights. This setup also better simulates the evolution of the

magnetic field under the influence of a realistic bottom boundary, since the depth of the box is much larger than the length scale of the flux tube.

Synthetic observations calculated from realistic simulation data provide a sufficient testing ground for the success of flow recovery methods. The following sections present how both the FLCT and DV methodologies have used numerical simulations of sunspots and to determine their effectiveness in photospheric ARs.

2.2 Methods

2.2.1 Data

For this study, two numerical simulations were used. The first being the R2D2 simulation presented by Hotta & Iijima (2020), which was used predominantly for training DV. The version of DV trained on this simulation will be dubbed $\mathrm{DV_{R2D2}}$ throughout this thesis. This run of the R2D2 code models the evolution of a twisted force-free flux tube placed, at a depth of 35Mm below the photosphere, under the influence of magnetoconvection. The simulation box has dimensions 98×98 Mm^2 with a periodic boundary in the horizontal direction and uniform grid spacing of $96\mathrm{kmpix}^{-1}$. In the vertical direction, the simulation covers a depth of 200Mm spaced over a non-uniform grid, and height up to 700km in the lower photosphere. The depth of this simulation allows the evolution of the flux tube to remain minimally affected by the bottom boundary. The R2D2 code uses a set of radiative MHD equations which, in this setup, assumes an ideal plasma (no physical resistivity) and no background magnetic field, so the flux tube evolves solely under the effects of magnetoconvection. Snapshots from the simulation have a temporal resolution of 120s over 180 hours of real-time simulation, this provides 37 hours of time before the emergence of the magnetic flux and over 100 hours of the decay process from the peak of the magnetic flux. Flow processes, therefore have a long time to evolve naturally with the rising magnetic flux prior to the emergence of the AR. Hence, the R2D2 simulation provides a realistic model of the interaction of convective plasmas and the magnetic field. An example intensity map and magnetogram, obtained from

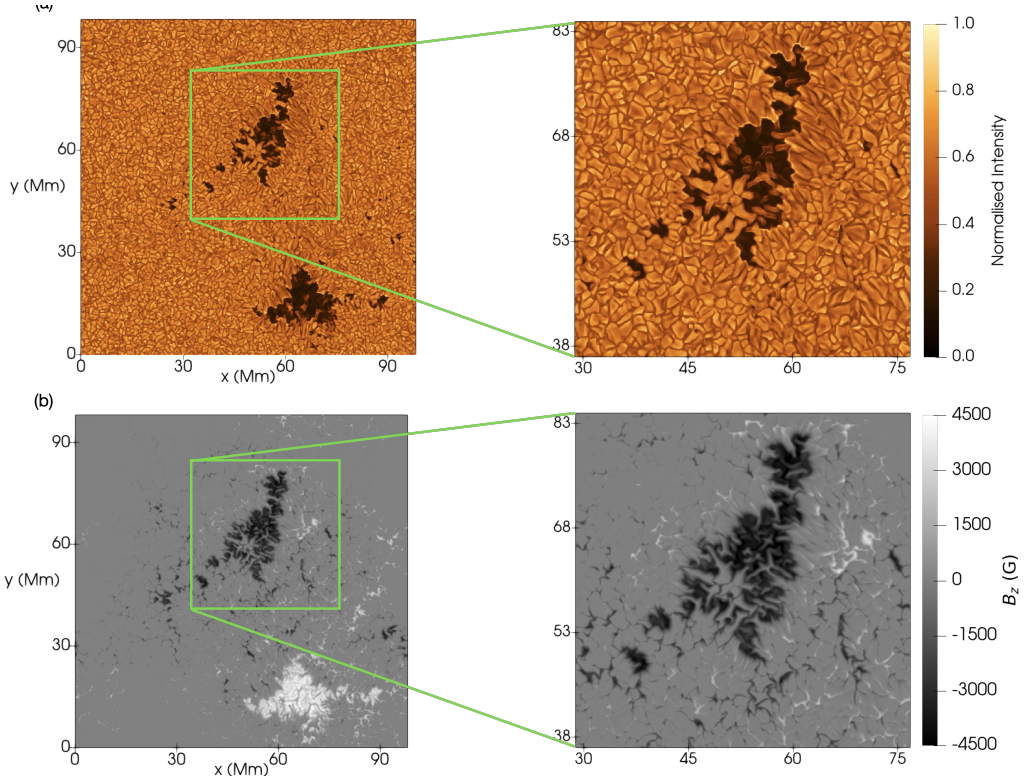

Figure 2.1: *An example of the fully developed AR at the time of peak magnetic flux (i.e. 60 hours after the start of simulation), from the R2D2 magnetoconvection simulation. On the top row, two panels show the white light intensity and the bottom row shows the corresponding magnetograms. The right-hand side shows a close-up view of the intensity and magnetogram of the region bounded by the green square on the left. The intensity was normalised by the maximum intensity of the entire region.*

the R2D2 simulation at a time of 60 hours after the start of the simulation (i.e. the time of peak magnetic flux in the photosphere), are shown in fig. 2.1.

The second dataset that was used was from the MURaM AR simulation presented in Rempel & Cheung (2014). This was used for the testing of flow recovery by the FLCT and the $\mathrm{DV}_{\mathrm{R2D2}}$ network. This simulation presents the realistic emergence of an untwisted semi-torus of magnetic flux in a box that has a periodic horizontal boundary with dimensions $147.5 \times 73.7 \mathrm{Mm}^2$. The depth of this box is 16.4Mm and shares the same $96 \mathrm{kmpix}^{-1}$ horizontal spatial resolution as the R2D2 simulation. Despite the shallow box relative to R2D2, this simulation includes the effects of turbulent diffusion, thus providing a more realistic thermal evolution of the plasma. An example frame of continuum intensity and magnetogram from

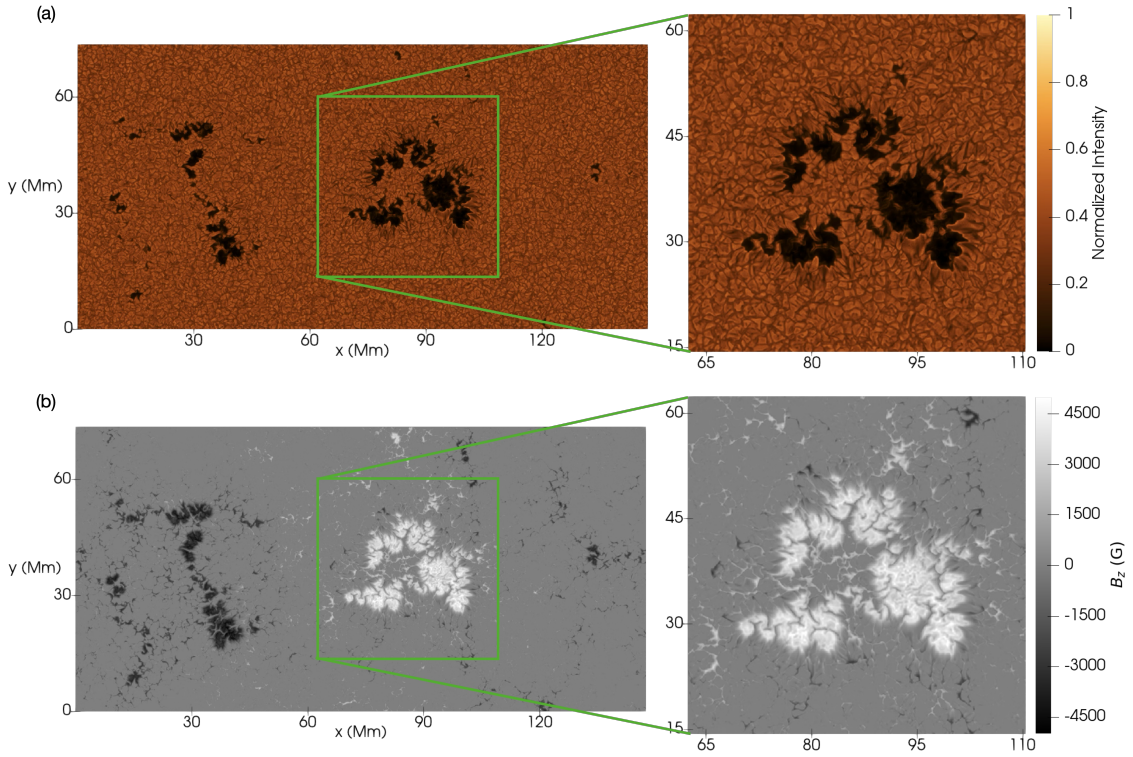

Figure 2.2: *An example slice of the AR at 83 hours after the start of the simulation, from the MURaM simulation. The panels in the left-hand column show the continuum intensity (a) and the magnetogram (b). The right-hand side panels show a close-up view of the intensity (top panel) and magnetogram (bottom panel) of the region bounded by the green square. The intensity was normalised between to be on a scale between 0 and 1.*

the simulation are presented in fig. 2.2.

The R2D2 and MURaM simulations offer valuable insight into different crucial aspects of the interactions of plasma flow and magnetic flux throughout AR evolution. Thus these provide an opportunity to properly test the ability of DV to accurately learn and reproduce flow features across simulations, from sub- to mesogranular scales, in the presence of intense magnetic flux.

The radiative transfer (RT) schemes between R2D2 and MURaM differ slightly. In R2D2, the RT scheme uses the Rosseland mean opacity and calculates values along only vertical rays by using a linear interpolation of the logarithm of the source function. The MURaM

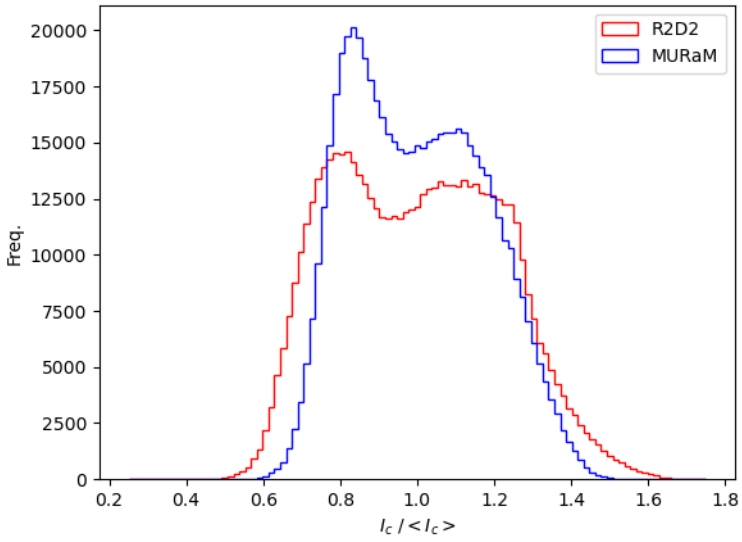

Figure 2.3: *The distribution of normalised intensity from the R2D2 (red) and MURaM (blue) simulations.*

simulation also uses the Rosseland mean opacity but uses a linear interpolation of the source function in calculations along rays in multiple directions. For full details of the schemes, we refer to the literature for R2D2 (Hotta & Iijima 2020) and MURaM (Vögler, A. et al. 2005). These differences in numerical schemes leads to differences in the intensities produced in the synthetic images. More specifically, since the R2D2 simulation uses a single ray, there is no RT calculated across the surface. Thus the contrast in the intensities will be greater, however the normalised intensities reveal a similar distribution to that of MURaM, see fig. 2.3. Testing DV on different simulations with variations in their numerical schemes, setups and included physics will provide insight into the limitations of DV when only trained on one simulation.

2.2.2 Preparing FLCT

It was found that images should be smoothed spatially to reduce the influence of small scales on the recovered velocities, and outputs should be averaged temporally in order to produce the best results. Previous works have found that FLCT displays the best performance over scales $>$ 2.5Mm, typically averaged over time windows of 30mins (e.g., see Rieutord et al.

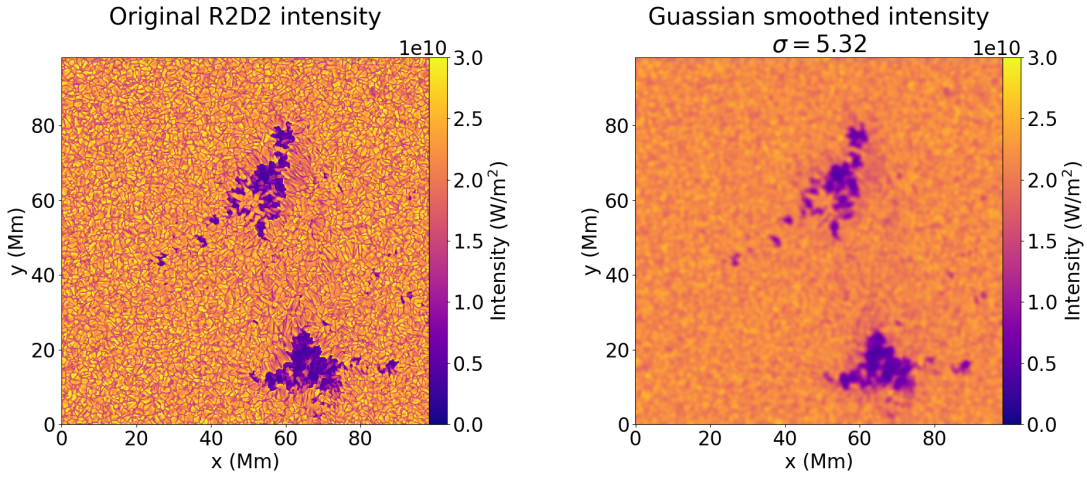

Figure 2.4: *Comparison of the synthesised continuum intensity from the R2D2 simulation with the same image with a Gaussian blur applied, typically used for the application of LCT methods. The value of σ in corresponds to a FWHM of $\approx$ 1200km, as was found to be the best sampling size by Verma et al. (2013). Note that only granular- and above-sized structures remain after the sampling.*

2001). For tracking granules, data should be taken from series with cadence no greater than 4mins (see Verma et al. 2013) as in this time the topology of the photosphere has changed significantly, not due to motions of convective cells, but due to their diffusion, in other words it becomes difficult to track motions of objects that have fragmented or emerged. Large errors are associated with the LCT method, for example velocities are largely underestimated by a factor of $\approx 1/3$ and do not necessarily have strong correlation with the ground truth data used for comparison from simulations (Simon et al. 1995, Verma et al. 2013).

An example of an image from the R2D2 simulation is shown in fig. 2.4, where the image has been blurred by a Gaussian of FWHM $\approx$ 1200km, which was found to provide the best performance by Verma, M. & Denker, C. (2011). The resulting image therefore contains features on the size of granules.

However, in the absence of any method like DV, and with the work by Tremblay et al. (2018) already highlighting a comparison of temporally and spatially averaged velocity fields, the comparison presented here uses the raw inputs at the same cadence and resolution that

was used to train DV. Thus, here the full capability of DV is shown with the available data and the improvement in performance over other methods can be detailed fully.

After testing parameters, it was found that for the R2D2 simulation with $\Delta t = 120$s, the optimum size of the apodising window was found to have a FWHM of $\approx$ 2500km and proved to have the highest correlation with features at mesogranular scales ($\approx$ 1800km) and up. The sampling window, in this case, is roughly double the size of that which is best for data with half the cadence (e.g., see Verma, M. & Denker, C. 2011, Verma et al. 2013). This was somewhat expected as its features will travel roughly twice the distance over double the time interval. This is also in alignment with what was found by Tremblay et al. (2018), where it was identified that FLCT performed the best at mesogranular scales.

Figure 2.4 shows an example of the resulting image used by FLCT after applying a high-pass filter. By convolving the image with a Gaussian to remove small-scales, the contrast of features on the desired scale are enhanced. This boosts the correlation of flows produced by FLCT with the target velocities. In this figure, a Gaussian with FWHM of $\approx$ 1200km was applied to a frame of the R2D2 simulation and shown to reveal features on granular scales and above.

It was highlighted by Simon et al. (1995) that velocity fields should be temporally averaged to smooth out the rather high associated errors with individual frames. However, such averaging forces the velocities to fit a normal distribution, which is not typically the true distribution of photospheric velocities, especially in the presence of ARs (see figs. 2.10 and 2.11). The resulting velocity fields produced by FLCT when images are smoothed and velocity fields are averaged are presented in fig. 2.5. Comparing the velocities displayed in panels (b) and (c), we see that without averaging the velocities are noisy and are only weakly correlated with the simulation whereas the averaged case reveals that FLCT velocities are more strongly correlated.

In fig. 2.6 scatter plots of the single and averaged recovered speeds are shown. It is clear from (a) that the flow speeds are largely underestimated and contains a lot of noise as expected from the study by Verma et al. (2013). In panel (b) the flow speeds are significantly

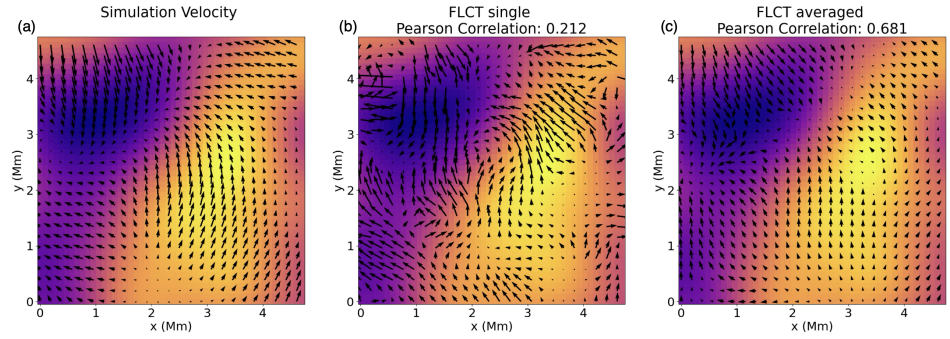

Figure 2.5: *Velocity fields from (a) simulation, (b) the velocity field recovered from simulation using FLCT on 2 consecutive frames and (c) the average of 15 recovered velocity fields. The background is coloured by the Gaussian-smoothed (*FWHM $= 1200$km*) continuum intensity of the simulation.*

better correlated with the simulation speeds, however the FLCT speeds remain noisy. As anticipated, the temporal smoothing via averaging does reduce the error in the velocity fields. This temporal averaging should be dealt with carefully as taking averages over longer windows does not continue to improve correlation with the true flows. As the time window increases, the velocity distribution tends toward a Gaussian centred about zero. Given that over a certain time period granules do not persist, the flow at a given point will be a random sample from the normal distribution and thus the error will increase as the averaged flow becomes more randomly distributed.

Additionally, when making like-for-like comparisons with alternative methods, which can provide accurate velocity fields with no post-processing, many of the small-scale features are lost. Hence, for this study, the FWHM of the apodising window is chosen to provide the most consistent results (FWHM $\approx$ 2500km) as measured by the Pearson correlation coefficient. No additional spatial smoothing or averaging is applied to the inputs to preserve the most information at all scales and provide a benchmark for comparisons with DV.

2.2.3 Training DV

For the training process, 3000 time-consecutive pairs of $4.8 \times 4.8\text{Mm}^2$ ($50 \times 50\text{pix}^2$) continuum intensity snapshots, from R2D2, were extracted with a time separation of the simulations

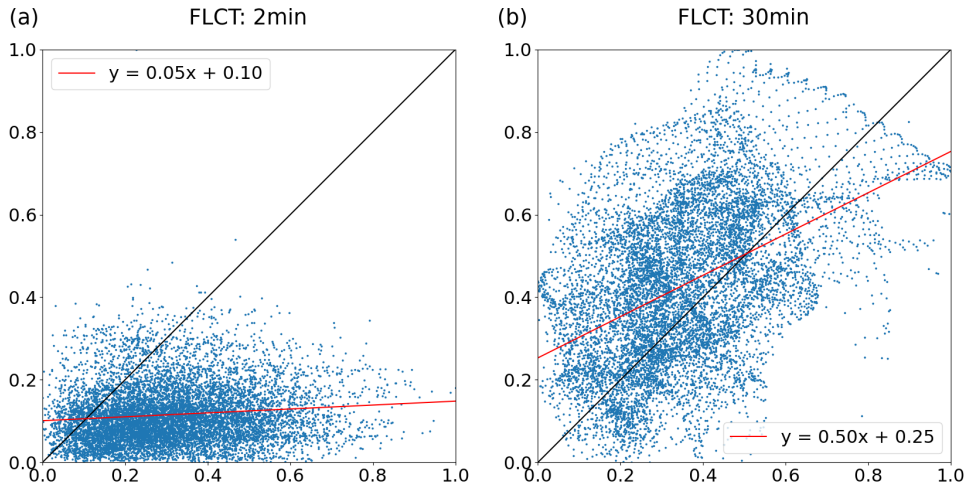

Figure 2.6: *Scatter plots showing how time-averaging of velocity fields produced by FLCT improves correlation between the normalised recovered and true speed of plasma. The curves of best fit for each distribution is shown in red and the ideal case is shown in black. Panel (a) shows the result of recovering only a single frame. Panel (b) show the 30 minute-averaged speeds.*

cadence (120s), along with their corresponding horizontal velocity fields. These training pairs were chosen at purely random times and positions across the entirety of the $\tau = 1$ surface of the simulation, and represent $< 1\%$ of the total data available. We suggest that this 120s cadence is suitable for testing the identification and analysis of flows at the (sub)granular scale since granules have a lifetime of $\approx$ 10min (e.g., see Bahng & Schwarzschild 1961). Whilst this is a relatively low cadence in comparison to some of the state-of-the-art telescope data available, it is a highly composite number, so most telescope imagery can be used at this lower cadence by selecting frames at this time apart. These pairs were filtered to ensure that there were no overlaps in the training data to avoid overfitting, that is to help ensure that DV is able to classify velocities well for new data just as well as the training set.

The training pairs were split into sets of containing 2000, 500 and 500 samples labelled as training, validation and testing, respectively. The training and validation set were used for training network. The network parameters are only updated if the predictive performance improves on the validation set, which is checked at the end of each training epoch, thus helping

to circumvent the problem of overfitting. The testing set was used to double check that the network performs well by hand and is presented in tables 2.1 and 2.3, and performance on the training data is omitted as this does not provide a true measure of how well DV performs on new data. However, the main focus is on testing with the MURaM data, thus this testing is presented in full as it is more indicative of the performance of the network acting on new data with different BCs, as would be typical of observations. Given the random nature of the sample selection, the snapshots should be representative of the flows present in all scenarios in the simulation.

It should be noted that, as with any supervised learning method, this training process poses the risk of introducing overfitting. If the number of samples from the training simulation is too large, the NN may learn only the images and flows from the training data and therefore be unable to predict the velocities accurately for new images. Because of this, we kept the proportion of training samples small to reduce this risk and reduce the computational cost of training the network.

The network was trained for 100 epochs (i.e. the network weights were adjusted over the entire set of training samples 100 times) using the Adam optimiser and the mean-squared error (MSE) loss defined as

$$\mathrm{MSE} = \frac{\sum_{i=0}^{n} \left|\mathbf{v}_{\mathrm{simulation,i}} - \mathbf{v}_{\mathrm{recovered,i}}\right|^2}{n}, \tag{2.1}$$

where i represents the observation (pixel) where the velocity is produced and n is the total number of observations. The Adam optimiser was chosen for efficiency as, by design, it includes a momentum for network weights and biases, updating the learning rate to allow for larger changes to rarely updated parameters and smaller changes for rapidly converging features—which also improves the stability of the network. Other key network hyperparameters include the number of filters, the kernel size, the number of residual blocks, the noise level, batch size and learning rate. In coalignment with the literature, which showed good results and efficiency, the chosen values for the list of hyperparameters were 64 filters, a kernel size of 3, 20 residual blocks, a noise multiplier of 0.001 and a learning rate of 10^{-4} (see the

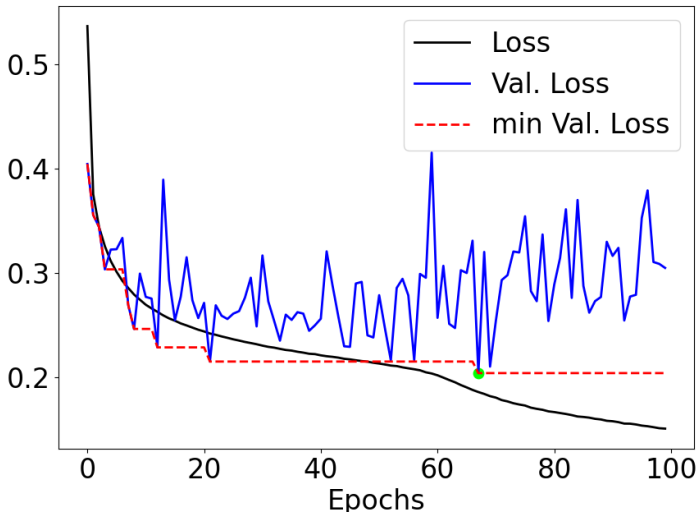

Figure 2.7: *Loss during the first 100 epochs of training from the DV during training. The black curve represents the loss value from the training data at each epoch and blue curve represents the loss value for the validation set (calculated at the end of each epoch). The red-dotted curve shows the minimum validation loss for the network at each epoch and the green point shows the minimum validation loss over the whole training process.*

network parameters chosen by Asensio Ramos, A. et al. 2017, Tremblay et al. 2018). The entire 100 epochs of training took approximately 100 minutes using NVIDIA V100 GPU and 10 hours on an Apple M2 card with 10 GPU cores.

The loss from the network throughout the training process is shown in fig. 2.7. This figure highlights that the network converges quickly and stably within the first few epochs of training, reaching close to its minimum value after around just 20 epochs. The network was allowed to continue training in which the learning stabilised at around 50 epochs before the validation loss started to increase despite the training loss beginning to decrease more rapidly around 60 epochs, suggesting that the network is now overfitting to the training samples, however the network reached its optimum performance at the 67th epoch.

In order to avoid overfitting, DV employs a number of methods. It does not update the network parameters unless the validation loss decreases, therefore the NN has an inbuilt feature allowing it to only update the weights if the generalisation of the network improves. It is seen that the tradeoff between number of epochs and the performance of the network

drops off after just a few iterations and thus training can be stopped early to improve the efficiency of the network. Additionally, during the training process, noise is added to the data by adding a set multiplier of the standard deviation in order to add variation to the data and avoid overfitting to the raw training samples. Frequent batch normalisation layers have also been used throughout the network in order to regularise the weights. Outside of the network architecture, the training set is augmented to include rotations and reflections of the training data so that the volume of training samples is increased without simply repeating any data. Furthermore, only small spatial windows were chosen in order to avoid the network seeing too much of the training simulation and help force it to understand flows in completely random regions.

2.2.4 Testing

Pairs of velocity fields and continuum images were used from the testing set in order to measure the performance of both FLCT and $\mathrm{DV}_{\mathrm{R2D2}}$. Examples of full frames from simulation are displayed in the results section also.

In the case of supervised learning, it is expected that NNs will perform well at reproducing information from data similar to what the NN has already seen during training. In other words, if the input data provided to the NN has similar properties to that of the training set, then the outputs produced will be consistent with the training set and hence consistent with the input data if it shares the same flow properties/physics of the training set. If the inputs have different properties to that of the training inputs, the outputted flows will be inconsistent with the training set flows, and most likely inconsistent with the test set flows as the NN has not learned to emulate the test simulation. In the case of providing new input data, the NN expects that most, if not all, of the properties of the input data will remain within the limits of the training set. Thus, if we test our model with images from a new simulation, which shares similar physics and properties of the training set, the outputs are expected to be reasonably good and will typically produce results that are in accordance with properties from the training set with some expected error. In our case, the physics and

properties of the MURaM and R2D2 simulations (e.g. cadence, resolution) are in agreement with each other. Other aspects of the simulation differ, the physical influence of a shallow bottom boundary, turbulent diffusion and an untwisted flux tube with a different initial shape present a somewhat different scenario to the R2D2 simulations.

Synthetic images from MURaM also differ slightly to R2D2 due to differences in the RT schemes. Thus, in order to ensure that the results produced by $\mathrm{DV}_{\mathrm{R2D2}}$ were consistent, the continuum images from MURaM had their histograms matched to those from the R2D2 simulation. This process has little impact on the actual structure of the images and simply ensures that the intensities are within the limits of what the network has been trained on. However, it should be noted that this makes some underlying assumptions about the images. Firstly, that all structures present in the simulations are the same, for example that granules, penumbral and umbral regions all share the same contrasts and topologies. Additionally, this implies that the physics presented are the same between simulations, e.g. the boundary and initial conditions—including field strength/magnetic flux density, twisting of the field and inclination of the field—play little-to-no role in altering the intensity of features. In the two simulations presented in this work in particular, these assumptions are not true, the radiative transfer is computed differently, the boundary conditions differ and the initial setup of the flux tube is different. However, most simulations will differ in subtle ways like these and no simulation can account for the full physics present in the Sun. Thus, histogram matching is a necessary step for optimising the use of DV with new data and this histogram matching is cheap and efficient for applying to new data and thus provides and opportunity to better understand how these differences will generate error. These same histogram-matched images were used as the inputs for FLCT to keep the preprocessing fair for both methods.

After the completion of the training process, individual velocity fields for full 1Mpix images took under 1 second to recover on both devices using DV. Ignoring the time taken for training, this is a 40× speedup compared with FLCT; the benefit of the speedup outweighs the time taken for training when considering large time series.

2.3 Results

Here we present the key results of our analysis for the success of DV in recovering velocities from the MURaM simulation when trained on flows from R2D2. In particular we present a number of metrics which show the success of DV in recovering flows at subgranular scales (96km) over active and quiet regions.

2.3.1 Metrics

In this section the effectiveness of DV at recovering AR flows is presented using a number of statistical metrics. The following list of metrics have been used:

- Pearson Correlation Coefficient (PC)

$$\mathrm{PC} = \frac{\sum(v_{\mathrm{MURaM}} - \bar{v}_{\mathrm{MURaM}})(v_{\mathrm{recovered}} - \bar{v}_{\mathrm{recovered}})}{\sqrt{\sum(v_{\mathrm{MURaM}} - \bar{v}_{\mathrm{MURaM}})^2 \sum(v_{\mathrm{recovered}} - \bar{v}_{\mathrm{recovered}})^2}}, \tag{2.2}$$

  where the bar over a variable represents the mean value. This correlation coefficient provides a measurement of the strength of the linear relationship between two variables between -1 and 1, where -1 describes a perfect negative correlation and positive 1 describes a perfect positive correlation. For perfectly recovered velocities we expect $PC = 1$, a strong correlation is typically identified as $0.5 < |PC| < 1$.

- Kolmogorov-Smirnov test (KS test)

$$\mathrm{KS} = \sup_x |\mathrm{CDF}_{\mathrm{MURaM}}(x) - \mathrm{CDF}_{\mathrm{recovered}}(x)|, \tag{2.3}$$

  in other words the supremum of the distance between the empirical cumulative distribution functions (CDFs) for the velocities from MURaM and DV, where sup defines the supremum function over the points x of the CDF. The test-statistic is interpreted as the maximum percent difference in vertical between values in the distribution (Hodges 1958).

- Root-mean-squared Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{\sum_{i=0}^{n}(v_{\text{MURaM},i} - v_{\text{recovered},i})^2}{n}}, \tag{2.4}$$

where n represents the total number pixels in the image. The RMSE provides an average of distance of the predicted values from the true values

- Median Relative Error (MRE)

$$\text{MRE} = \text{Median}\left(\frac{|v_{\text{MURaM}} - v_{\text{recovered}}|}{|v_{\text{MURaM}}|}\right). \tag{2.5}$$

The relative error shows the difference between the predicted and true values as a proportion of the true value, i.e. percent error. Since the velocities can be typically large, in any location where the velocity may be, in actuality, close to 0 and is predicted to be large will cause the relative error to explode. For this reason, the median is chosen as opposed to the mean as this will ignore the extreme values.

- Normalised dot product (NDP)

$$\text{Normalised dot product} = \frac{\mathbf{v}_{\text{MURaM}} \cdot \mathbf{v}_{\text{recovered}}}{|\mathbf{v}_{\text{MURaM}}||\mathbf{v}_{\text{recovered}}|}. \tag{2.6}$$

The normalised dot product is the cosine of the angle between the predicted and true velocity vectors. This shows the alignment of the recovered vectors as a value between -1 and 1 where 1 represents a scenario where all the predicted vectors are parallel to the true vectors, -1 are antiparallel and 0 represents a case where all the vectors are at right angles to the true vectors.

2.3.2 Comparing FLCT and DV performance with MURaM

Two-dimensional velocity fields were recovered from the MURaM simulation by applying DV to time-consecutive pairs of continuum intensity, separated by a cadence of 2 minutes. The x- and y- components of the velocity fields were estimated for $\tau = 1$ at various times throughout

the evolution of the photosphere. At $t \approx 2$h the magnetic flux in the photosphere has a mean value of < 1G and a maximum value of 20G. Therefore, this highlights an example of when the simulation presents mostly QS dynamics in the photosphere, seen in fig. 2.8. An example of a time when intense magnetic flux densities are present in an AR is presented at $t \approx 83$h. These particular times were chosen to highlight the capability of the flow recovery methods since they provide examples of the full range of flow dynamics in the presence of QS and AR levels of magnetic flux density. Examples of the original and recovered velocity fields are shown in fig. 2.8 and fig. 2.9. The recovered velocity fields were compared to the original simulated velocities.

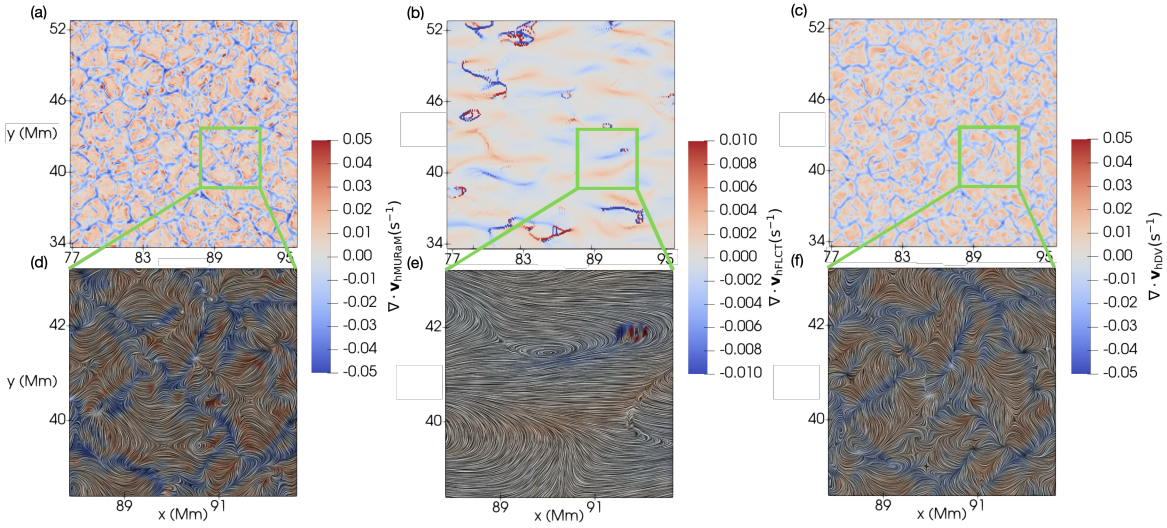

Figure 2.8: *A close view of the divergence of the horizontal components of velocity, 2 hours into the MURaM simulation. This figure highlights a region where* $>$ kG *magnetic flux emerges later in the simulation. Panel (a) shows the divergence field from simulation over a zoomed-in section of the region highlighted in figure 2.2. Panel (d) shows a closeup of the LIC with the velocity field superimposed on the divergence field of the highlighted region from panel (a). Panels (b) and (c) show the same regions as panels (a), where the velocities have been estimated by FLCT and DV, respectively. Panels (e) and (f) show the same region as (a), again using velocities from FLCT and DV, respectively.*

Figure 2.8 (a) shows the divergence of the horizontal velocity field, $\nabla \cdot \mathbf{v}_h$, from the MURaM simulation at $t = 2$ hours, when the magnetic flux rope is still being convected below the surface. Panels (b) and (c) present the divergence of the horizontal velocity field recovered by FLCT and DV, respectively, over the same region as panel (a). Panels (d), (e)

and (f) show the line integral convolution (LIC; created by convolving a white noise filter with the velocity field and integrating over the field lines) over a zoomed-in region, distinguished by the green square. LIC was first presented by Cabral & Leedom (1993) and it emphasizes the streamlines of the horizontal velocity field using dark lines. The zoomed-in region highlights a region where complex magnetic structures later appear. The red and blue colouration represents regions of positive and negative horizontal divergence, which is referred to just as the divergence. Panels (b) and (e) highlight that FLCT is unable to recover any meaningful flow features when being used for instantaneous recovery, with no discernible similarities with the simulation flows. From panel (c), $\mathrm{DV_{R2D2}}$ is able to reproduce the location of divergent features in the flow apparently well, with the magnitude of the divergences not being as high as those present in the simulation. This discrepancy in the divergence is likely due to differences between speeds in the R2D2 and MURaM simulations. $\mathrm{DV_{R2D2}}$ is able to recover diverging features accurately from MURaM at length scales of intergranular lanes (< 500km).

Figure 2.9 presents the same region as was shown in fig. 2.8, much later in the simulation ($t = 83$ hours) after the magnetic flux tube has emerged. Panels (a) and (d) once again show the divergence and flows from simulation, which now highlight a region at the edge of a pore, where many intricate flow structures are present. In this instance, some of the strongest diverging features have been identified by FLCT and rough large scale shapes can be discerned such as the boundary of the pore and large granules that have been distorted by the strong horizontal inflows/outflows. However, flows on local scales are still highly inconsistent with those from the simulation. On the other hand, $\mathrm{DV_{R2D2}}$ presents the ability to distinguish flow behaviour between that over the pore and that on the edge in panels (c) and (f). The distinguishing difference between the simulated and NN-recovered flows is that $\mathrm{DV_{R2D2}}$ is not able to identify the exact topology of small-scale vortices (< 100km) and saddle flows within the pore. However, DV is able to to identify some of the apparent swirling motions, despite velocities having a magnitude of $\approx 1/10$ that of the flows surrounding the AR. This is to be somewhat expected as the images exhibit far less contrast in these regions.

Figures 2.10, 2.11, 2.13 and 2.14 give further insight into the results discussed above;

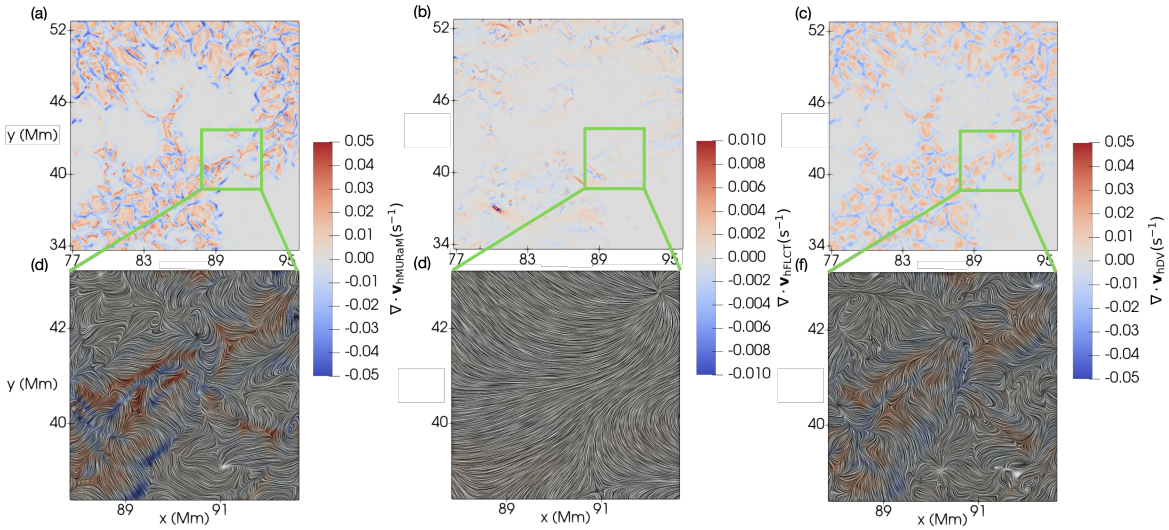

Figure 2.9: *The same visualization as Figure 2.8 is shown, at a time after the AR has emerged (83 hours). The zoomed-in regions highlight the boundary of a pore, containing complex flow structures such as saddles and vortices.*

they present the distributions of different velocity field components from simulation, FLCT and DV_{R2D2} and the scattering of these quantities, respectively as a means of analysing the correlation and trends of the recovered velocity fields.

Figure 2.10 (a) and (b) display the horizontal components of the simulation and FLCT recovered velocity fields. The recovered velocities follow a normal distribution, similar to the simulation, however their standard deviations are significantly lower with velocities being concentrated much more strongly about the mean of 0 leading to significant underestimation of values. The Gaussian width is slightly wider in the x direction, suggesting that FLCT favours producing velocities in one direction over another, when the frame is rotated this pattern remains consistent suggesting a bug in the pyflct package used. The underestimation is apparent in panel (c), which highlights that the speed over quiet regions has a mean that is about 1/3 that of the true mean. In addition to this, the velocities are not distributed well spatially, by the divergence where all information is lost due to the nature of FLCT velocities matching a simple normal distribution.

Table 2.2 displays some values relating to the shape of the simulated and recovered velocity field distributions, further highlighting that the FLCT distributions contain large outliers and

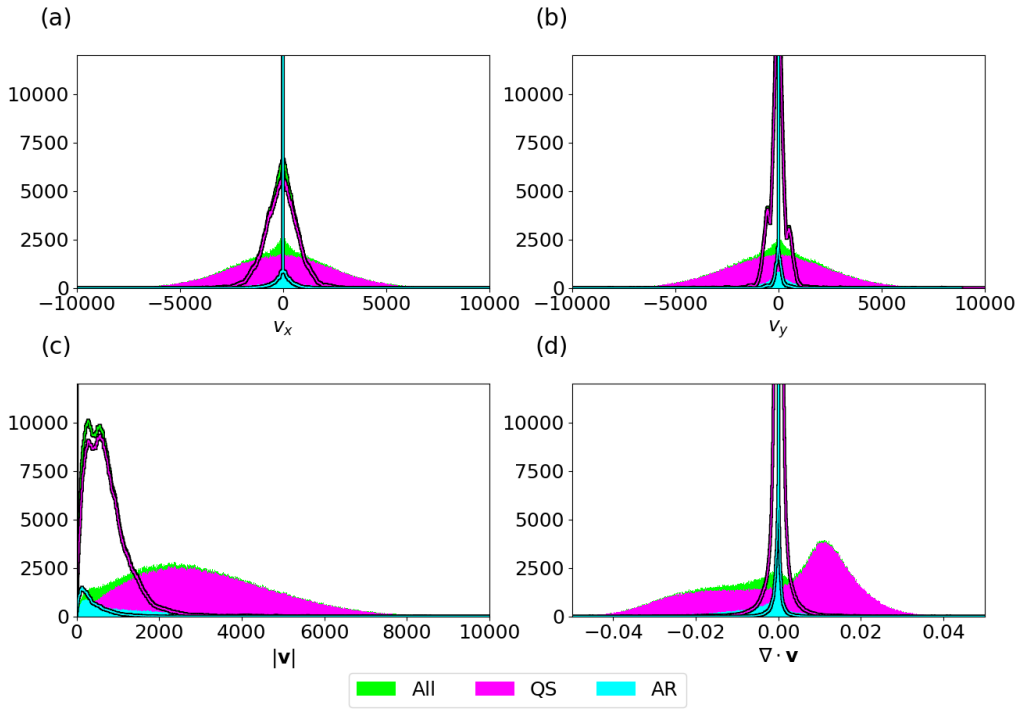

Figure 2.10: *Histograms showing the distribution of different components of the velocity field at a time of 83 hours, for the original simulation (solid coloured backgrounds) and the flows recovered by FLCT (the outlined curves) over regions of $>$ kG magnetic flux (blue), $<$ kG flux (magenta) and over the entire field of view (green). Panels (a) and (b) show the distributions for the x and y components of velocity. Panel (c) shows the distribution of speed and (d) shows the distribution of divergence. All are shown in cgs units. Note the stark contrast in shapes of the distributions from FLCT compared to the simulation which indicate the method is unable to replicate the target distribution which will lead to large errors in the velocity fields.*

do not closely resemble those from simulation. Table 2.4 quantifies the performance of the FLCT recovery, showing that all components and derivatives are poorly correlated with the simulation in both active and quiet regions. Further, the distributions differ by up to 50% as measured by the KS test and flows are generally unaligned with an average angle between the FLCT-recovered and simulation flows of $\approx 75°$. The significant difference between the distributions, described by the KS test value, are shown in fig. 2.12, where the centring about 0 creates a steep gradient in the CDFs from FLCT. The scatter plots in fig. 2.13 also reveal how the velocities do not correlate well with the simulation flows. The lines of best

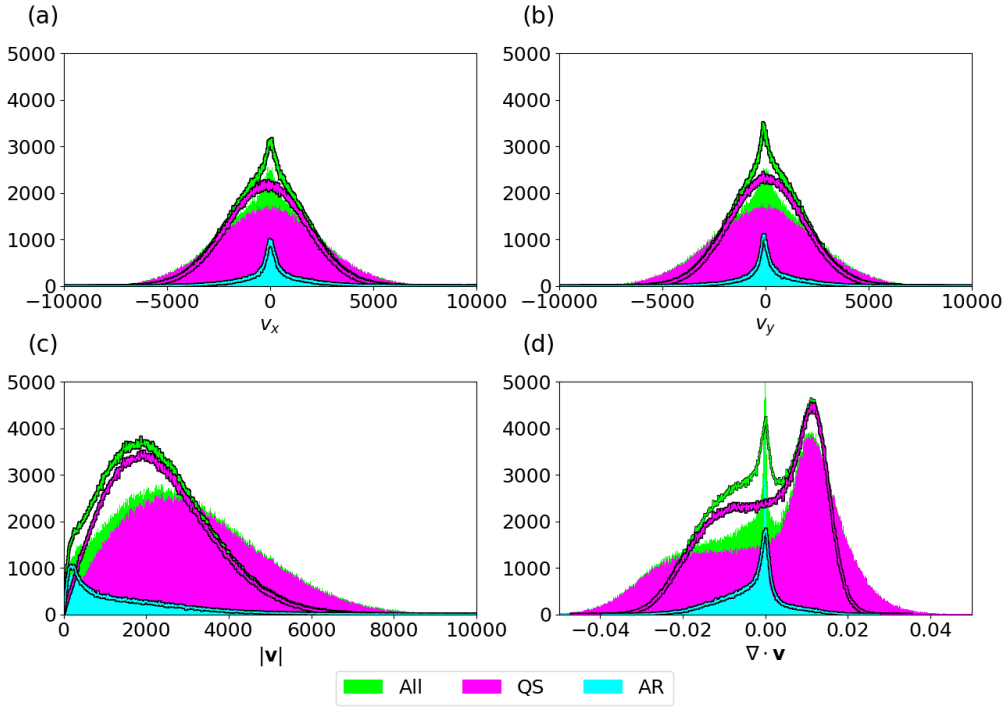

Figure 2.11: *Same as fig. 2.10 except using recovered velocities from* $\mathrm{DV}_{\mathrm{R2D2}}$. *Note here the much closer shapes in distributions compared to the previous figure, indicating that the performance from DV is a significant improvement over FLCT.*

fit show that the horizontal velocity components actually closely follow the same trend as the simulation, but this is simply an artifact of extreme values that are produced in some instances skewing the distribution.

The histograms in fig. 2.11 reveal that in both active and quiet regions $\mathrm{DV}_{\mathrm{R2D2}}$ is able to match the horizontal velocity distributions from MURaM well. It appears that DV typically underestimates the velocities present in QS regions, but the distributions still match the shapes and general properties of the flows well. This is backed by a Kolmogorov-Smirnov test, which highlights that the components of the horizontal velocity vector are chosen from a distribution differing by a maximum of $\approx 10\%$ from the true distribution (see table 2.2 and fig. 2.12). This deviation is due to the tendency of DV to underestimate the velocities shown in panels (a) and (b), respectively. The distribution of speeds shown in panel (c) and better highlights the underestimation in velocity components, combining to show an overall

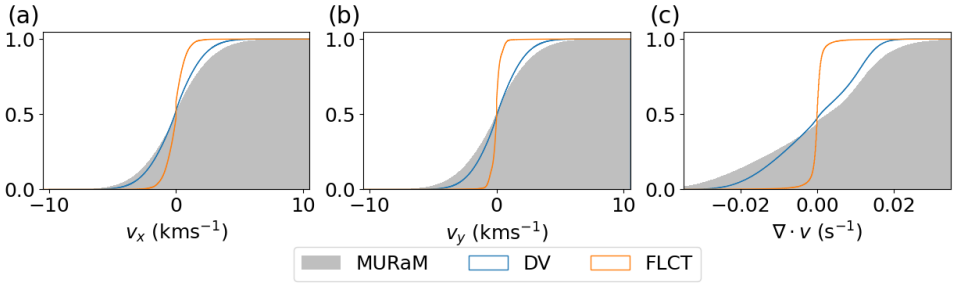

Figure 2.12: *The cumulative distribution function for the horizontal velocity components (a) v_x and (b) v_y, and the (c) divergence of the 2D velocity field $\nabla \cdot \mathbf{v}$. The CDF for the MURaM velocities are shown as the filled-in grey background, and the CDFs for DV and FLCT are shown in blue and orange, respectively. Note that the CDFs from DV only differs a small amount from MURaM's, except in divergence where the gradient is relatively steeper, whereas FLCT produces CDFs that are significantly steeper and differ a lot from the target distribution.*

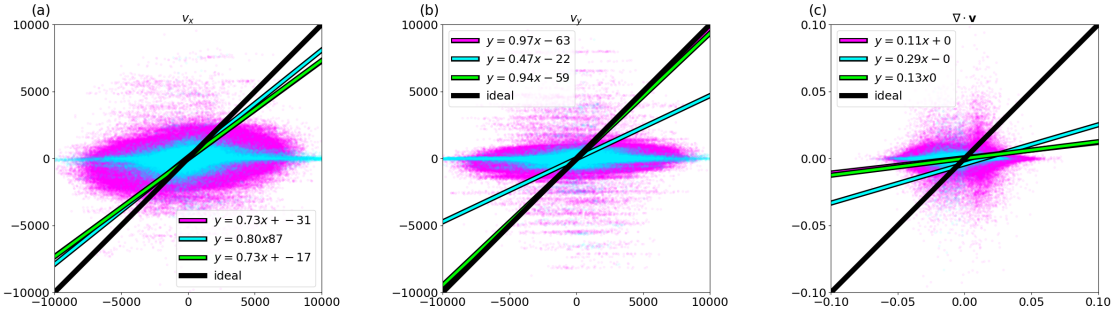

Figure 2.13: *Scatter plots of components of velocities (a) and (b) and their divergence (c) recovered from by FLCT and their lines of best fit, plotted against the simulated values from MURaM, shown in cgs units. Values over ARs are shown in blue, QS are shown in purple. The green line shows the overall line of best fit over all types of regions and the black line shows the target line of best fit ($y = x$). Points have had their opacity lowered to remove outliers and highlight where the strongest spread of velocities lie.*

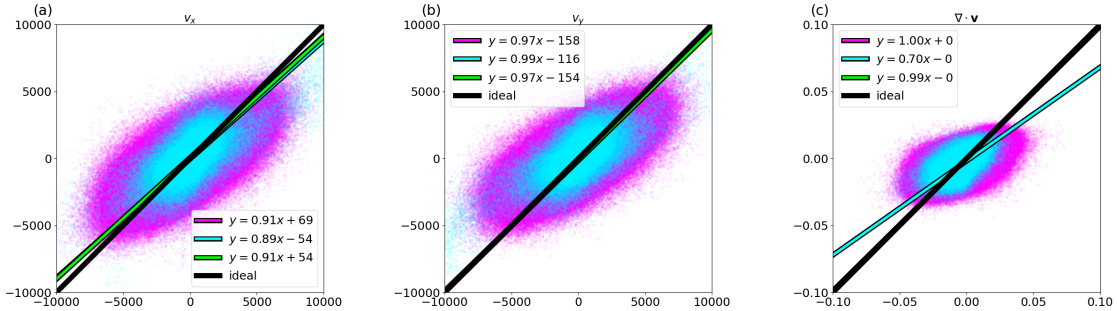

Figure 2.14: *Same as fig. 2.14, however using comparing velocities predicted by* $\mathrm{DV}_{\mathrm{R2D2}}$ *against the simulation velocities.*

reduced speed, implying that errors in predicting velocities largely come from a shift in the component distributions. The speeds are still within the same order of magnitude despite this, highlighted by table 2.2. The divergence distribution of the DV velocity field also shares similar properties to the simulation, capturing the bimodal nature of the distribution, which is result of the mostly non-diverging nature of the flow within the AR and the positively divergent granules that dominate the QS-like regions. The key properties of the divergence match well despite the differences in the variance and errors introduced by taking derivatives of data.

The scatter plots in fig. 2.14 highlight the strong correlation of the velocities over all regions, and that flow speeds are typically slightly underestimated. The divergence is more weakly correlated over the AR, and carries more error than individual components of velocity (as to be expected from taking derivatives of data containing errors), however over the QS the overall divergences of the recovered velocity fields are highly correlated. In fact, it is seen that DV performs particularly well at reproducing the distributions of AR flows, only struggling to capture the exact values of divergence in the AR, shown in panel (c) and fig. 2.11.

Tables 2.1 and 2.2 show that $\mathrm{DV}_{\mathrm{R2D2}}$ is able to replicate the properties of the velocity fields well about the means of the distributions, seen by the means and interquartile range (IQR) matching well for all presented values. Only the divergence and curl, where the error is expected to be increased from taking derivatives of estimated data, present issues. Despite this, the divergences of recovered flows correlate strongly in QS regions. Flows still correlate

well, but not as strongly, within ARs where the contrast in images becomes much weaker and the dynamics differ largely due to the suppression of plasma motions by the magnetic field. Comparing the values in the two tables, it can be seen that performance on the test data of the training simulation is similar to that of the MURaM test simulation. DV does appear to more closely match the distributions (range and IQR) for R2D2, but this is to be expected as all the simulation properties fit within the parameters already seen by DV. In comparison with the distributions of properties over AR, table 2.4 suggests some potentially contradictory results and that flows identified in these regions by $\mathrm{DV_{R2D2}}$ are more weakly correlated and that flows contain significant errors. However, the table presents values that are spatially dependent whereas the distributions are not. In other words recovered flows in ARs are not recovered spatially as well as those in QS regions, but the properties of the flows in these regions align better with the simulation.

In order to quantify the success of the recreated flows, tables 2.3 and 2.4 displays metrics describing the accuracy of flows from $\mathrm{DV_{R2D2}}$, including their alignment and associated errors when compared against the simulation flows. The square of the PC describes the explained variance captured by the method. The recovered flows, as mentioned above, are highly correlated with the simulated flows. This high correlation shows that a significant amount of information in the flow is recovered. The PC of the divergence is only slightly impacted over QS regions despite taking derivatives of velocities, which will invariably include some error. The correlation of the divergence within the pore becomes more moderate compared to the surrounding flows but still shows a statistically significant correlation to the original flows. We also observe the same trend in the alignment of the flows measured by the NDP, that is the cosine of the angle between the recovered flows and the simulation flows. A normalised dot product close to 1 would imply the flow is well aligned, and a value close to zero represents a flow with no alignment with the target. The $\mathrm{DV_{R2D2}}$-recovered flows have a mean alignment of around 0.674 over quiet regions, and a lesser alignment over the AR of around 0.501. The inverse cosine of these values corresponds to mean differences in angles of $\approx 47^\circ$ over the QS-like regions and $\approx 60^\circ$ over ARs. This is a significant increase over

the averaged difference of $\approx 75^\circ$ in the alignment of the FLCT-recovered flows. Additionally, the median normalised dot product sits at 0.871 and 0.774 for the NN-recovered flows over both regions, thus half of the velocity vectors are within 30° and 40° of their true direction over their respective regions. Furthermore, the RMSE and the MRE are shown. The RMSE provides an average distance of the NN velocities from the true velocities, however this can be easily skewed where large velocities are involved. The MRE is more representative of the performance of the network, as it is less skewed by outliers in the recovered flows. Since the velocities and divergences from $\mathrm{DV}_{\mathrm{R2D2}}$ correlate well with the simulation and the flows are aligned well over the majority of the simulated surface, the apparently high relative error should be taken in context. Velocities produced by DV have less variance than the velocities in MURaM, but correlate strongly. Therefore the topology and features of the flow field may remain consistent with the simulation but show a different magnitude which produces a consistently large relative error due to the speed of the flows. This difference in magnitude is highlighted by fig. 2.11 (c), where there is an underestimation in the speed of flows, particularly over the QS regions. These values are consistent with the mean relative error shown by Tremblay et al. (2018), which still prove stronger than other flow recovery methods. The values from the table pertaining to the performance of $\mathrm{DV}_{\mathrm{R2D2}}$ on the R2D2 simulation (see table 2.3) behave very similarly to $\mathrm{DV}_{\mathrm{R2D2}}$ on the MURaM simulation, with the differences being that the network more closely matches the distributions from the R2D2 simulation when compared with MURaM (as measured by the KS test) and the performance is enhanced over the ARs. This enhanced performance over the AR, versus MURaM, is likely due to the fact that the images from MURaM were histogram matched to match the properties of the QS from R2D2, thus limiting $\mathrm{DV}_{\mathrm{R2D2}}$'s capacity to correctly reproduce flows within the AR. The fact that the in NN performance across the simulations implies that the network has generalised well and shows promise that it will perform similarly across other new datasets with similar properties.

		Minimum	Maximum	Mean	IQR
v_x (kms^{-1})	R2D2	-10.0	10.0	14.8×10^{-2}	3.14
	DV	-12.2	14.2	-88.8×10^{-3}	3.19
v_y (kms^{-1})	R2D2	-10.0	10.0	-52.5×10^{-3}	3.18
	DV	-12.9	11.7	0.119	2.95
$\nabla \cdot \mathbf{v}_h$ (s^{-1})	R2D2	-0.0861	0.0635	8.04×10^{-7}	0.0185
	DV	0.0660	0.0409	1.07×10^{-5}	0.0185
$(\nabla \times \mathbf{v}_h)_z$ (s^{-1})	R2D2	-0.0886	0.0905	4.32×10^{-8}	0.0503
	DV	-0.0679	0.0696	-1.25×10^{-7}	0.0381

Table 2.1: *Table displaying the quantitative properties of the distributions of the horizontal velocity fields ($\mathbf{v}_h$) from the R2D2 simulation and those recovered from* $\mathrm{DV}_{\mathrm{R2D2}}$.

		Minimum	Maximum	Mean	IQR
v_x (kms^{-1})	MURaM	-11.1	12.4	-62.1×10^{-3}	3.11
	DV	-10.2	12.5	-175×10^{-3}	2.83
	FLCT	-285	125	-61.5×10^{-3}	0.810
v_y (kms^{-1})	MURaM	-13.5	12.5	-88.3×10^{-3}	3.13
	DV	-12.2	11.0	119×10^{-3}	2.65
	FLCT	-37.4	28.6	-31.4×10^{-3}	0.312
$\nabla \cdot \mathbf{v}_h$ (s^{-1})	MURaM	-0.0815	-0.0728	5.52×10^{-7}	0.0238
	DV	-0.0521	0.0355	$8.22 \times 10^{-}6$	0.0197
	FLCT	-1.49	1.50	-1.60×10^{-6}	1.05×10^{-3}
$(\nabla \times \mathbf{v}_h)_z$ (s^{-1})	MURaM	-0.0958	0.114	5.61×10^{-7}	8.63×10^{-3}
	DV	-0.0614	0.0662	-3.61×10^{-5}	6.55×10^{-3}
	FLCT	-1.50	1.50	-1.33×10^{-6}	6.41×10^{-4}

Table 2.2: *Table displaying quantitive properties of the distributions of the horizontal velocity fields ($\mathbf{v}_h$) from the MURaM simulation and those recovered from FLCT and* $\mathrm{DV}_{\mathrm{R2D2}}$.

	PC		KS test statistic		RMSE		MRE		NDP	
	QS	AR	QS	AR	QS	AR	QS	AR	QS	AR
$\mathbf{v}$	-	-	-	-	-	-	-	-	0.689	0.605
v_x	0.715	0.799	0.0201	0.0279	1.85	1.45	0.607	0.756	-	-
v_y	0.701	0.804	0.0395	0.0378	1.86	1.50	0.684	0.816	-	-
$\nabla \cdot \mathbf{v}_h$	0.485	0.589	0.0233	0.0702	0.014	0.00797	0.594	0.784	-	-
$(\nabla \times \mathbf{v}_h)_z$	0.248	0.508	0.0396	0.438	0.010	0.072	1.40	1.44	-	-

Table 2.3: *Metrics for the horizontal velocity fields recovered by* $\mathrm{DV_{R2D2}}$ *from the testing dataset of the R2D2 simulation. The values are split over QS conditions (*$<$ kG*) and AR conditions (*$>$ kG*). The RMSE for velocities and derivatives is given in units of* $\mathrm{kms^{-1}}$ *and* $\mathrm{s^{-1}}$*, respectively.*

		PC		KS test statistic		RMSE		MRE		NDP	
		QS	AR	QS	AR	QS	AR	QS	AR	QS	AR
$\mathbf{v}$	DV	-	-	-	-	-	-	-	-	0.674	0.501
	FLCT	-	-	-	-	-	-	-	-	0.254	0.249
v_x	DV	0.750	0.712	0.072	0.054	1.69	1.47	0.710	0.886	-	-
	FLCT	0.256	0.167	0.494	0.258	2.43	2.00	3.48	81.9	-	-
v_y	DV	0.752	0.717	0.096	0.039	1.68	1.49	0.745	0.950	-	-
	FLCT	0.186	0.0723	0.496	0.254	2.47	2.09	8.62	204	-	-
$\nabla \cdot \mathbf{v}_h$	DV	0.723	0.549	0.112	0.052	0.0116	0.00843	0.642	0.903	-	-
	FLCT	0.0267	0.0589	0.506	0.348	0.0172	0.0102	20.7	47.1	-	-
$(\nabla \times \mathbf{v}_h)_z$	DV	0.335	0.444	0.053	0.059	0.0118	0.0113	1.51	1.34	-	-
	FLCT	0.00293	0.00223	0.458	0.297	0.0125	0.0102	7.51	49.2	-	-

Table 2.4: *Metrics comparing the recovered horizontal velocity field from FLCT and* $\mathrm{DV_{R2D2}}$ *to the original velocities from the MURaM simulation. Velocities were taken over the entire* $\tau = 1$ *surface at a time where the* $>$ kG *magnetic flux was present in the photosphere (i.e 83 hours after the beginning of the simulation). Compared values are split over QS conditions (*$<$ kG*) and AR conditions (*$>$ kG*). The RMSE for velocities and derivatives is given in units of* $\mathrm{kms^{-1}}$ *and* $\mathrm{s^{-1}}$*, respectively.*

2.4 Summary

For the first time, DV was validated on the recovery plasma flow features on scales from the mesogranular (several Mm) to subgranular (0.096Mm) scales in simulated AR flows by cross-testing with a second simulation. This was done in order to remove any bias that may have come from DV overfitting to the training set. In addition to this, we compared the flow recovery to the instantaneous outputs from FLCT, to benchmark the like-for-like effectiveness of the method with a widely-used method. With these we determined how effective the methods are, especially in QS regions and ARs where magnetic flux density is higher and therefore strongly influences plasma flow.

One of the main things found was that DV can recover topological flow features accurately, including diverging and swirling motions at granular scales, within the most complex regions of an AR. Because of this, DV can be used to uncover information about horizontal inflow and outflow in ARs at scales that haven't been studied much before. Little topological information was present in the instantaneous velocity fields from FLCT that was present in the simulations. It is suspected that with preprocessing of data, i.e. performing a high-pass filter, and averaging the results, FLCT would be able to produce velocity fields that are more consistent with the simulation, as determined by Tremblay et al. (2018), and therefore provide detailed topological information at mesogranular scales, but not below.

Considering the performance of $\mathrm{DV}_{\mathrm{R2D2}}$ on a second simulation, it is expected that it will also perform well on observational data, when the training set is treated to mimic properties of telescope imagery. It should be noted that before applying a version of DV to observational data, there are a number of features present in images such as aberrations due to the point spread function (PSF) of telescopes, noise due to turbulence in the Earth's atmosphere (when dealing with ground-based telescopes) and other physics in the Sun, which has not been modelled by simulations. Each of these should be studied carefully in order to understand their impact on the flow recovery by DV. Some of these have been highlighted in appendices A and B. Whilst there are algorithms for reducing the effects of atmospheric noise such as

blind deconvolution (e.g., see Liu et al. 2022), there also exists a number of high-resolution space-based telescopes which remain unaffected by atmospheric effects. Some high-resolution datasets could be readily used with our model and it is expected that results would be fairly consistent with the ground truth, however the actual ground truth data does not exist, thus tests with these are meaningless. With FLCT however, these issues are not as concerning as the flows are identified by optically tracking features and the proper motions of these features would be unaffected by consistent effects such as the lensing caused by a telescope. In addition, the typical smoothing of images and averaging velocity fields reduces randomly distributed errors, presenting an advantage over DV for working with observations containing large (mesogranular) scale features of interest.

Convolving images with the optics of the Sunrise/IMAX instrument was shown to have only a small influence on DVs accuracy, reducing the PC of velocity fields by ≈ 0.1 when applied to both MURaM and R2D2 (see appendix A). The training set used for the NN is mutable however, so images may be convolved with the optics of any instrument that monitors a similar region of the photosphere, and then DV be retrained for application using the modified images. It is also possible, where the point spread function of an instrument is known, to deconvolve optical effects from observations, this would make our trained model, $\mathrm{DV}_{\mathrm{R2D2}}$ suitable for recovering the velocity fields from deconvolved images so long as their cadence and resolution match closely with the R2D2 training set.

Overall, $\mathrm{DV}_{\mathrm{R2D2}}$ performed well, though it struggled when recovering smaller-scale features like vortices in umbral regions. This is likely because strong magnetic fields in those areas suppress plasma motions, which only account for a small amount of the data in the training set. Still, $\mathrm{DV}_{\mathrm{R2D2}}$ was able to differentiate between various flow behaviours, showing it could adapt to different scenarios within the photosphere. $\mathrm{DV}_{\mathrm{R2D2}}$ effectively reproduced the velocity field distributions, though there were some small inaccuracies in the tails of the velocity distributions. Inaccuracies such as these are somewhat expected, as supervised learning methods with continuous outputs will tend to produce values that are smoothly distributed about typical values, observed in the training set. We measured $\mathrm{DV}_{\mathrm{R2D2}}$'s ac-

curacy using Pearson correlation, RMSE, and alignment (dot product) on singular outputs. In QS regions, the results were good, but in ARs, the performance was a bit lower. For example, the mean alignment in QS regions was 0.674, but it dropped to 0.501 in ARs. This is also likely due to the impact that magnetic flux has on plasma motion as well as the inputs having a lower contrast in umbral regions as opposed to the typical contrasts present in QS regions. Additionally, the lower correlation for divergence and curl in ARs suggests that DV has a harder time recovering more complex flow structures when intense magnetic fields are involved. These issues likely arise due to differences in how intensity maps are constructed between training and testing data, which becomes more noticeable in ARs where magnetic fields drastically change plasma flow. Our results highlight that despite differences in BCs in synthetic data and the exact numerical setup, we are still able to produce consistent results that provide insight into the types of features present over ARs, even if their exact topology is not captured fully.

As presented in this chapter, DV is a robust tool, which can be trained to handle complex dynamical scenarios in a solar context. The validation that was performed shows that DV has the capability to recover the smallest scale flows present in synthesised images from simulations, instantaneously with a good level of accuracy, vastly outperforming FLCT in both speed and accuracy. In conclusion, this study shows that DV has a lot of potential for recovering velocity fields from continuum intensity data including ARs. While it does have some limitations within umbral regions, DV is a promising tool for studying solar photospheric flows and provides a solid foundation for future applications. Future work could focus on reducing the discrepancies in AR regions, particularly in recovering smaller-scale features and better capturing how magnetic flux affects plasma flows.

In the next chapter, the ability for DeepVel to recover Lagrangian coherent structures in flows is tested as a further way to determine the capability of the network. These LCSs highlight the dynamics present in the flow and thus determine a method for understanding the flow physics that are captured by the network.

Chapter 3

Recovering Lagrangian Coherent Structures with DeepVel

3.1 Introduction

A crucial step towards understanding the Sun is to be able to recover physically meaningful flows from observations. Plasma flow is important in guiding the dynamics of the Sun including its magnetic field (see van Driel-Gesztelyi & Green 2015, Tziotziou et al. 2023, for some examples). In fact Alfvén's theorem (see Alfvén 1942) states that magnetic flux is conserved in ideal MHD conditions, in other words the magnetic field lines become frozen into the plasma and are advected by fluid motions. Despite the Sun not presenting ideal conditions, the magnetic field and plasma on the Sun do influence each others' structure, Hence, plasma has the ability to transport material and energy about the solar surface.

The transport of material by fluids can be pinpointed by distinguishing material surfaces within the plasma flow that act as transport barriers. These transport barriers prevent materials, such as magnetic flux, from passing through them. In the context of solar plasma, the magnetic field is typically strongly linked to the motions of the plasma in the photosphere where the magnetic Reynolds number is large (e.g., see Parker 1963). To this end, the transport of magnetic flux across the photosphere is dominated by advection through the plasma.

In the Lagrangian point of view, material surfaces are defined by tracing the trajectories of particles in the flow and identifying regions which are not crossed by them. By their definition of following particle trajectories, these persisting structures are Lagrangian in their nature and thus are Lagrangian coherent structures. Since fluid elements do not cross these structures, they partition the flow into regions that exhibit similar dynamics, thus defining a skeleton for the dynamics of the flow. Thus, LCS theory facilitates a novel and precise method for analysing flow properties by identifying the strong local repelling, attracting and shearing surfaces (Haller & Yuan 2000, Haller 2015*b*). In other words, the skeleton defined by LCSs decomposes flows into dynamic building blocks. An example of material surfaces/LCSs appearing in nature is given in fig. 3.1 where the material surfaces, that separate fluid layers in the atmosphere, are distinguishable by eye. In particular the spot will be delimited by a strong elliptical LCS and within this will be a complex network of attracting and repelling structures that contribute to the constant vorticity observed. One can loosely think of the material surfaces as ever-shifting walls which either pump fluid into or out from either side of it, guiding and restructuring the surrounding flow.

Analysing the ability of a method to recover LCSs can give a significant insight into the effectiveness of a method to predict the precise flow topology and hence the evolving dynamics of photospheric flows. They also provide a key insight into the transport of material by the flow, i.e. the influence that plasma flows have on transporting magnetic flux.

It has already been evidenced by Chian et al. (2014) that coherent structures dominate much of the photosphere, and the presence of attracting LCSs correspond to concentrations of magnetic flux in the photosphere. Additionally, Chian et al. (2019) highlighted how the skeleton of the flow, provided by the FTLE field, exposes the locations of the centres and boundaries of supergranular cells. Furthermore, Chian et al. (2020) observed the presence of persisting Lagrangian saddles, which indicate recurring vortical motions within a complex magnetic region. More recently, Silva et al. (2023) showed that prior to flux emergence, there are changes to the FTLE field at the mesogranular scale up to 10 hours before the emergence of an AR, indicating how Lagrangian analysis can be used to understand and even predict

Figure 3.1: *An observation of Jupiter's atmosphere, namely the great red spot, from the Voyager 1 mission in February 1979. Image courtesy of NASA/JPL and Björn Jónsson. There is a distinct stratification of the gas, separated by material surfaces which distinguish Lagrangian coherent structures such as an elliptical LCS surrounding the spot, acting as a transport barrier.*

the evolution of magnetic features at the solar surface.

As $\mathrm{DV_{R2D2}}$ (presented in chapter 2) has been established to be able to identify velocity fields from simulations well, it is of interest to determine to what extent the network is capable of producing physically meaningful flows by comparing coherent structures from simulated and predicted velocity fields. To this end, $\mathrm{DV_{R2D2}}$ was used to study topological plasma flow features throughout the total evolution of a simulated active region by tracking tracers and reproducing coherent patterns by means of calculating the FTLE fields. In the previous chapter, the success of the NNs predictive ability was determined in terms of correlations and error of velocity fields. As mentioned, advecting particles along a time series of the flow will compound errors that are contained within velocity fields. Thus, calculating LCSs allows the study of how well the flow topology has been identified correctly by $\mathrm{DV_{R2D2}}$ and hence support the understanding of how significantly the flow dynamics deviate from their true state throughout a time series of recovered velocity fields. Additionally, the calculation of FTLEs involves taking derivatives of the motions of particles in the flow, which are therefore expected to amplify any errors present in the velocity fields. Therefore, analysing what has been recovered by $\mathrm{DV_{R2D2}}$ will help to determine how well the predicted velocity fields emulate photospheric flow physics from the R2D2 simulation and further determine how well the model generalises to other realistic and complex flows present in the MURaM simulation.

3.2 Methods

As presented in chapter 2, a version of DV was trained on the R2D2 simulation from Hotta & Iijima (2020), using randomly selecting pairs of subimages and their corresponding velocity fields from the $\tau = 1$ optical surface, representing the photosphere. This NN ($\mathrm{DV_{R2D2}}$) is able to identify instantaneous velocity fields from photospheric simulations containing ARs, with a good accuracy. The flows produced by the network strongly correlate with the simulation and diverging structures are captured well.

In this work, the flow dynamics of the R2D2 and MURaM simulations were studied by seeking coherent flow structures by identifying transport barriers in the flow. These

transport barriers act as a skeleton for the flow, dictating the general flow topology and thus the movement of material within the flow. Repelling and attracting material surfaces were identified by means of identifying ridges in the FTLE field, calculated forward and backward in time (see, e.g. Haller 2015*b*). In areas where these ridges overlap, there may be complex flow behaviours present such as saddles and vortices. Regions with small FTLE values (i.e. that contain no ridges) highlight regions with simple flows. In order to calculate the FTLE field, a uniform grid of particles is advected over a time-dependent velocity field and the distance between neighbouring particles is measured. The full mathematical description of this process is already given in chapter 1. The integrations of the velocity fields were performed over two, different length, time intervals of 20 and 100 minutes. These times were selected as they reveal structures that correspond to granular and mesogranular scales, which have been shown to correlate with magnetic flux (Chian et al. 2014, Silva et al. 2023). Therefore, it is anticipated that being able to replicate LCSs and changes in them at these scales will determine the success of DV in emulating the physics of simulation.

3.3 Results

To further assess the capability of the recovered velocities from $\mathrm{DV}_{\mathrm{R2D2}}$, its ability to reproduce the natural transport barriers from, and therefore emulate the flow physics of, the MURaM and R2D2 simulations was tested. The forward- and backward-FTLE field was computed by integrating over short (20 minutes) times and long (100 minutes) times, thus recovering granular and mesogranular scales, respectively. This was performed at different key times in the simulations to reveal how flow structures change when high concentrations of magnetic flux are present in the photosphere. The times chosen reflect where the photosphere is mostly magnetically quiet and when it is magnetically active. Results concerning their nature and distributions are discussed below.

3.3.1 Coherent Structures in the R2D2 Simulation

These results present a best-case scenario for validating $\mathrm{DV}_{\mathrm{R2D2}}$'s ability to reproduce coherent flow structures, since the image properties, flow properties and physics of the simulation are all the same as the training set. Thus, the flows produced by the network correlate more highly with the R2D2 simulation flows than with the MURaM simulation flows, thus it is expected that more of the LCSs identified correspond strongly with those present in the simulation.

Figures 3.2 and 3.3 outline the ability of $\mathrm{DV}_{\mathrm{R2D2}}$ to capture flow dynamics by reproducing transport barriers at granular and mesogranular scales, respectively. The longer time integration used to reveal mesogranular flow structures highlights the presence of structures that persist longer. These mesogranular structures, in particular the attracting structures, appear more strongly than their granular scale counterpart, shown in fig. 3.3. These stronger structures indicate a greater attraction that persists over periods of hours during the peak of an ARs emergence. Almost all of the ridges present in the simulation are captured by $\mathrm{DV}_{\mathrm{R2D2}}$ at both of the featured scales, only underestimating the attracting ridges and overestimating the repelling ridges slightly. At the granular scale, both the attracting and repelling structures weaken over the pore, where flows are suppressed by the magnetic field, however the attracting structures show a lesser reduction in strength suggesting that plasma is flowing into the pore. $\mathrm{DV}_{\mathrm{R2D2}}$ is able to identify this difference, however the consistency of the topology of the flow structures with the simulation is reduced at the mesogranular scale, suggesting that flows within pores have highers errors that compound enough to generate notable difference at larger scales. This is expected as the FTLE is calculated using multiple velocity fields that contain errors, yet the enhanced attraction within the pore is still captured. The differences in the FTLE magnitude hints that $\mathrm{DV}_{\mathrm{R2D2}}$ is able to distinguish the difference between regions that are dominated by granulation and magnetic fields. The right-most column of the figure shows that in the R2D2 simulation, there is the presence coherent structures indicative of Evershed-like flows (e.g., see the discussion by Strecker, H. & Bello González, N. 2022) surrounding the AR. This is represented by the strong attracting ridge, which delimits the

pore boundary and is further delimited by a strong repelling ridge in both the short and long-term integrated fields. This defines a region with strong small and large scale horizontal flows into the AR, which is typically present around strong pores and sunspots. The NN correctly identifies the FTLE ridges delimiting the AR at both granular and mesogranular scales. These results suggest that the NN is able to emulate the flow physics within and around the AR well.

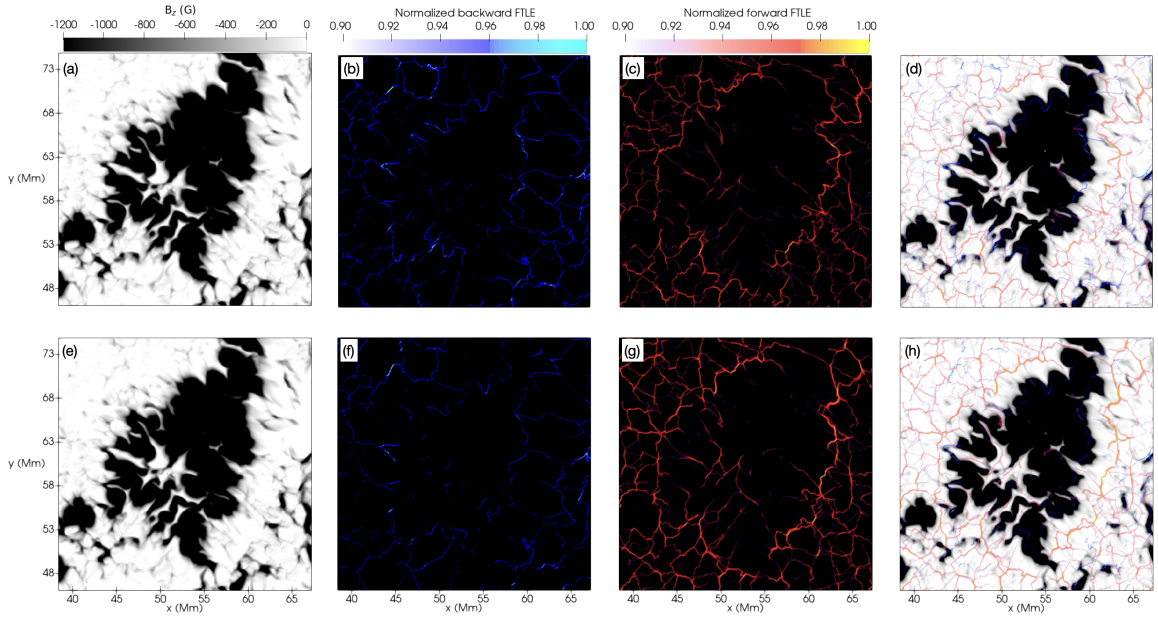

Figure 3.2: *The normalised forward- (red) and backward-FTLE (blue) field superimposed on the magnetic field, corresponds to the highlighted region in fig. 2.1 in the R2D2 simulation at the time of peak magnetic flux in the photosphere (60 hours). Panels (a) and (e) show the magnetic field averaged over the time of integration. (b) and (f) show the backward-FTLE ridges resulting from a 20-minute integration of the simulated and recovered velocity fields, respectively. Panels (c) and (g) show the resulting forward integrated FTLE fields. Panels (d) and (h) both the forward and backward integrated FTLE fields superimposed on the magnetic field. The ridges resulting from the 20-minute integration correspond to features on the length scales of granules.*

It is shown further in fig. 3.4 that the distributions of the forward-FTLE fields are very closely matched by $\mathrm{DV}_{\mathrm{R2D2}}$ with the target distributions from simulation. In fact a Kolmogorov-Smirnov test reveals a test statistic of ≈ 0 implying that there is no discernible difference between the FTLE distributions from the simulation or NN flows. The distributions are shown at times before the emergence of any magnetic flux, as well as during the

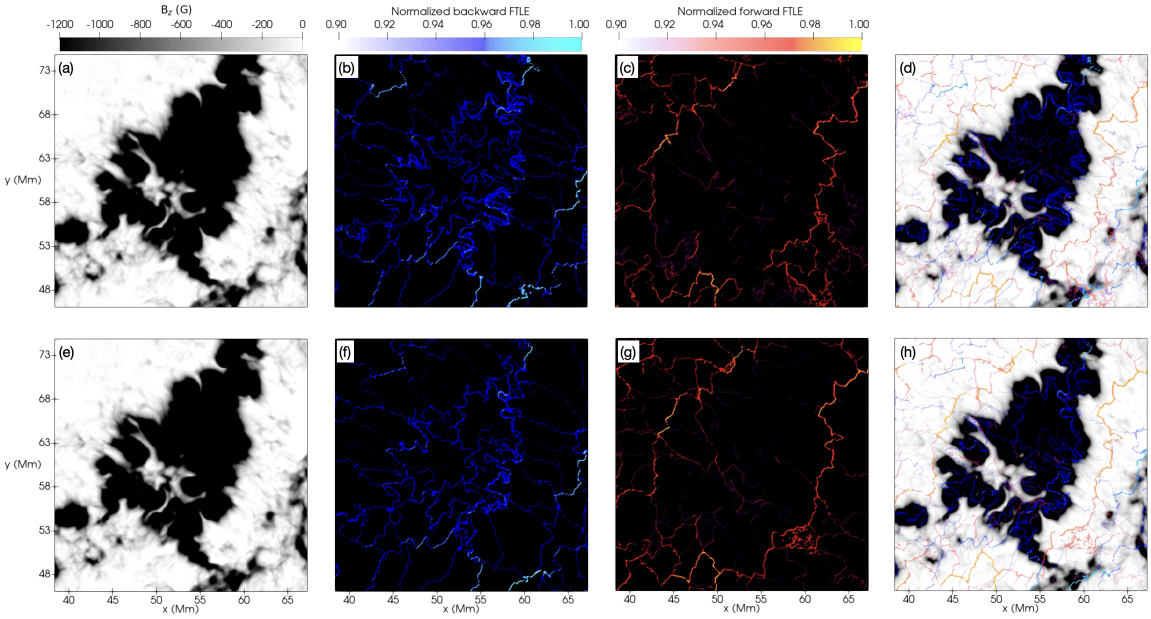

Figure 3.3: *The same as Figure 3.2 however the FTLE ridges here result from a 100-minute integration corresponding to mesogranular features.*

emergence and peak, and the decay of an AR. Throughout the emergence and evolution of the AR, up until the decay, the mean FTLE value over the $<$ kG shifts slightly to the left. This shift in the mean indicates a small global change in QS regions where little magnetic flux is present at all times. A similar trend is seen in the regions where $>$ kG magnetic flux becomes present, however a more significant change occurs as the magnetic flux emerges and reaches its peak density. This reduction in the mean forward-FTLE corresponds to reduction in repelling barriers over the AR, suggesting the presence of stronger attracting structures or a simplification of the flow behaviour in general. In all cases, $\mathrm{DV}_{\mathrm{R2D2}}$ is able to distinguish the slight differences in the FTLE field between regions where low concentrations and high concentrations of magnetic flux density are present. This result also highlights the possibility of monitoring changes, at granular scales, in the transport of magnetic material by the flow. This has implications for the potential use of forecasting and tracking of dense magnetic fields with just the flow field.

What these results strongly highlight is that $\mathrm{DV}_{\mathrm{R2D2}}$ has successfully learned to emulate the physics of the R2D2 simulation well.

3.3.2 Coherent Structures in the MURaM Simulation

The same process, as described above for the R2D2 simulation was carried out for the MURaM AR simulation (Rempel & Cheung 2014). As established, in chapter 2, the initial setups for these simulations differ. In this case an untwisted semi-torus of flux was advected through the bottom boundary, a depth of 16Mm below the surface, and then allowed to evolve under the influence of convective plasma motions. Therefore, the behaviour exhibited by the flows is expected to differ somewhat from the R2D2 simulation, in which a twisted cylindrical flux tube is inserted deep beneath the photosphere, at a depth of 35Mm beneath the surface.

In figs. 3.5 and 3.6, the short and long term integrated forward and backward FTLE fields are shown for the MURaM simulation as well as those derived from velocities predicted by $\mathrm{DV}_{\mathrm{R2D2}}$. Panels (b) and (c) of both figures highlight a reduction in strength of the FTLE ridges over the AR, similar to the ridges over the R2D2 pore where flows are suppressed. Furthermore, attracting barriers delimit the entire pore indicating the presence of inflows at the umbral boundary. The red repelling barriers are slightly weaker over the pore but still dominate much of the region, indicating a horizontally outward transport of material by the flow. These repelling barriers suggest that the magnetic field is dissipating by means of the plasma transporting it radially outward from the centre of this region. At the time shown in the figure, the simulated AR is in fact decaying and magnetic flux density is decreasing, thus the barriers correspond to the physical behaviour of the magnetic fields. $\mathrm{DV}_{\mathrm{R2D2}}$ is able to reconstruct the repelling structures very well, maintaining most of the fine structure at granular scales. $\mathrm{DV}_{\mathrm{R2D2}}$ is also able to reconstruct the strongest attracting features, however it largely overestimates much of the weaker ridges over the pore. Thus the NN is able to capture short-lived and granule-sized diverging flow features well, but less so the converging flow features on these scales. In the 100-minute integrated field, presented in fig. 3.6, the network is able to reconstruct both attracting and repelling barriers on mesogranular scales well also. In the MURaM simulation, the dynamics exhibited by the flow are different to the R2D2 simulation. Within the LCSs, revealed by the FTLE ridges, it is found that the Evershed-type flows present in the R2D2 AR are not present in the MURaM AR. The

Evershed-like flows are not present in the MURaM simulation due to insufficient BCs used for computational efficiency at the photospheric surface in the simulation (see Rempel & Cheung 2014), whereas some assemblance of penumbrae are found in the R2D2 simulation (see Hotta & Iijima 2020), however these are still less prominent than in observations. $\mathrm{DV}_{\mathrm{R2D2}}$ produces flows that are mostly consistent with the MURaM flows, and the flow dynamics of the simulation have been identified well as seen through the FTLE fields. Thus showing that the network has been able to generalise well to the MURaM simulation flows, since the NN does not simply mimic the flow features present in training set and is therefore emulating the flow physics present in the simulations. Therefore we determine that $\mathrm{DV}_{\mathrm{R2D2}}$ is able to consistently produce flow structures equally well across simulations. In particular the forward-FTLE fields correlate well with the originals at granular scales and up, whereas the backward-FTLE field, while still maintaining much of the structure, is underestimated at granular scales. This indicates that DV is able to generalise well to new data and, if trained to deal with the challenges present in observations, it would perform well at identifying flow dynamics on the surface of the actual Sun.

In fig. 3.7 the distribution of the 20-minute forward-FTLE field is shown for 2 different regions in the flow, over times when there is little magnetic flux present in the photosphere and when the AR is fully emerged (2 and 83 hours, respectively). Panels in the left column of the figure present the forward-FTLE distribution over a section of the photosphere where <kG magnetic flux is present throughout the simulation. On the other hand, the right column shows the distribution over the region where >kG magnetic flux has emerged in the later time. The distributions are shown in blue and red for early and late times, respectively. Panel (a) of the aforementioned figure shows that the distribution in quiet regions, at both early and late times, remains almost constant, and panel (b) reveals that there is a difference when magnetic flux emerges into the photosphere. We can see this more clearly as the mean over the AR changes significantly, relative to the range of the FTLEs ($\approx 10\%$ difference). DV is able to reproduce this change in the shape of the distribution, however not as accurately, due to the errors in the predicted velocities. In this instance, a Kolmogorov-Smirnov test produces

a test statistic of up to 0.06, in other words the maximum difference in the distributions is 6%. Thus, whilst LCSs are produced near-perfectly for the training simulation, the cross-test with MURaM reveals small errors, but measurable errors. Despite this, DV is recovering key physical aspects of flows at the granular scale, which determine the its behaviour in the presence of $>$ kG magnetic flux.

3.4 Summary

This study is the first to present an analysis on how flows recovered by DV are able to reproduce the transport barriers present in simulated flows. The version of the network used here presents a great performance in replicating the overall transport dynamics of photospheric plasma in regions dominated by intense concentrations of magnetic flux in the simulation. The reconstruction of attracting and repelling barriers matched well with simulations of ARs, especially at granular and mesogranular scales. This confirms that despite the inevitable errors in the recovered velocities, DV is able to reconstruct enough information such that these did not majorly impact particle motions and trajectories in the flow, which shows that DV is a robust tool for studying the transport of material at small-scales by plasma flows. Furthermore, this indicates that time-dependent structures do not significantly deviate from their true nature when identified by the NN. On top of this, $\mathrm{DV}_{\mathrm{R2D2}}$ successfully captured the distinct changes in flow behaviour during the formation of ARs, like when magnetic flux suppresses plasma motion or when magnetic fields are spread outward.

The ability of DV to emulate flow properties from the provided training set means that it should recover the characteristics of flows, so long as they are in line with flow features present in the training simulation. Typically, NNs may suffer from over-fitting and create features that are not present or fail to reproduce anything coherent in the worst case, when passing new data through the network. However, DV did not introduce the topological features that were present in the training data into the predicted flows for images from MURaM, where those topological features were not present. This is a strong indicator that the model constructed by DV from R2D2 is good at generalising to new data containing similar features

with differing dynamics and that DV is able to emulate the properties of AR flows well from the training set.

In addition to these points, the FTLE field exhibits a signal which is indicative of the presence of magnetic flux in the photosphere. Changes in LCSs produce a quantifiable difference in the distribution of repelling structures, this difference highlights what is similarly shown by Silva et al. (2023). Thus the presence of high concentrations of magnetic flux alters the structure of LCSs in the flow. Not only this, but DV shows that it is a suitable methodology for recovering enough information on granular and mesogranular scales to accurately identify such structures and changes within them.

The following chapter studies how changes in the flow topology may be used as a proxy for identifying where interesting magnetic features are present in the photosphere, by constructing a classifier for the presence of intense magnetic flux with a threshold for changes in the topology of LCSs.

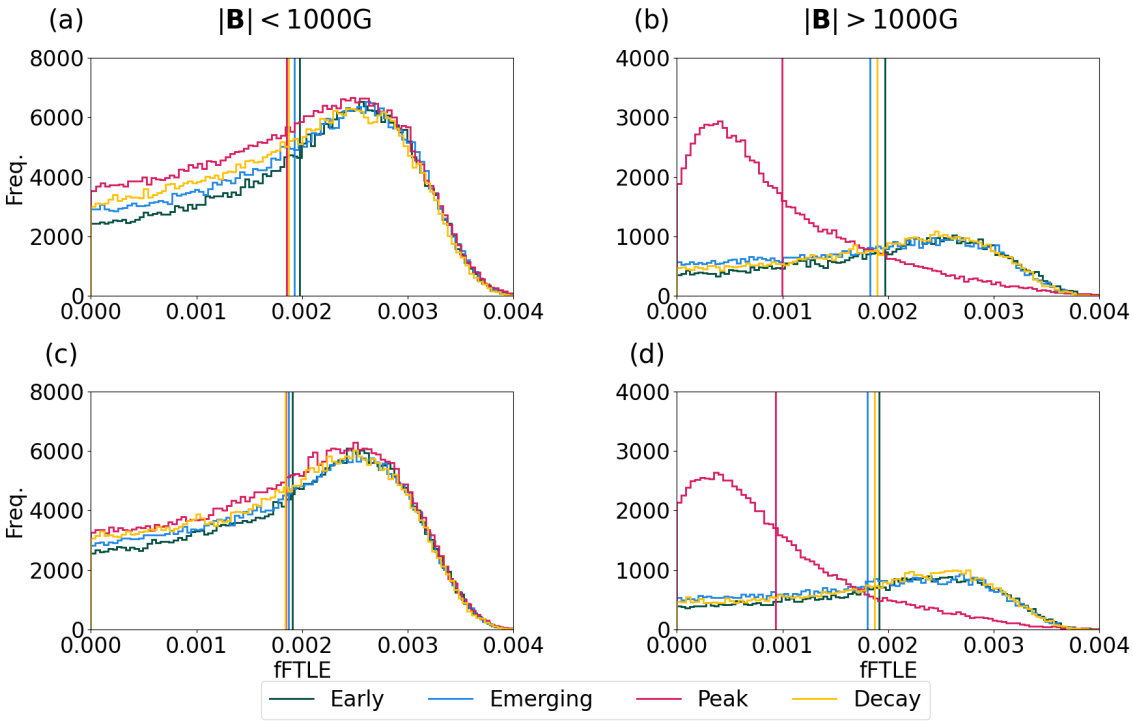

Figure 3.4: *Distributions of the 20-minute forward-FTLE field, at various times in the R2D2 simulation during the AR lifetime: 1h 40min (early, when there is little concentration of magnetic flux in the photosphere; dark green), 40h (during the emergence of the AR in the photopshere; blue), 60h (the time of peak magnetic flux density in the photopshere; magenta), 166h 40min (close to the time of total decay of the AR; yellow). The top row shows the FTLE distributions calculated from the simulation velocities and the bottom shows the distributions calculated using velocities predicted by* DV_{R2D2}*. The left column shows the distributions over a region which contains* $<$ kG *magnetic flux density throughout the evolution of the photospheric surface and the right column displays the distributions over a region where* $>$ kG *flux is present during the emergence and peak of the AR. The vertical lines show the corresponding positive foward-FTLE means for each distribution*

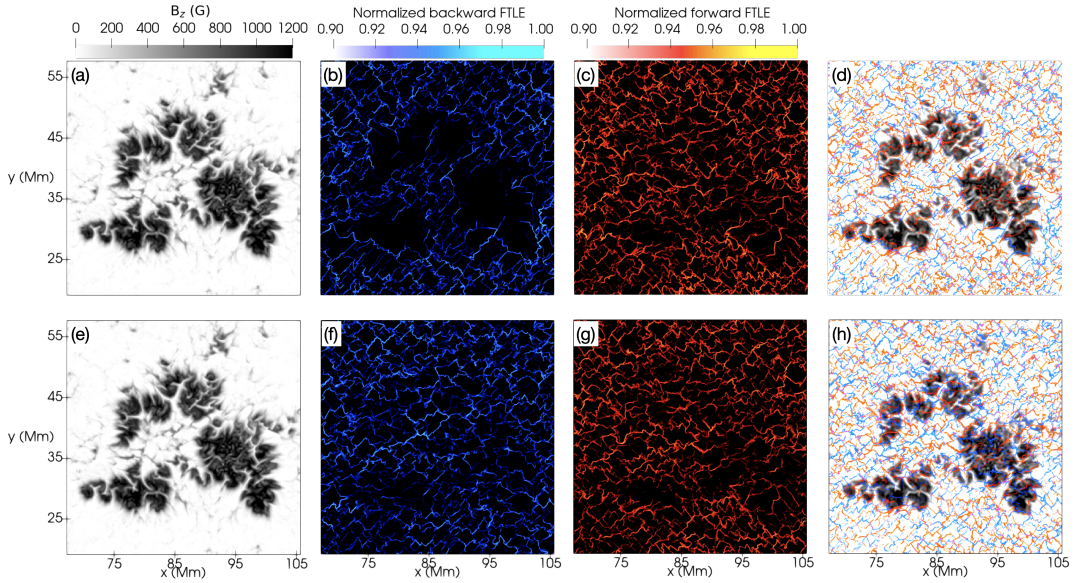

Figure 3.5: *The same as fig. 3.2, however highlighting the region from the MURaM simulation shown in fig. 2.2.*

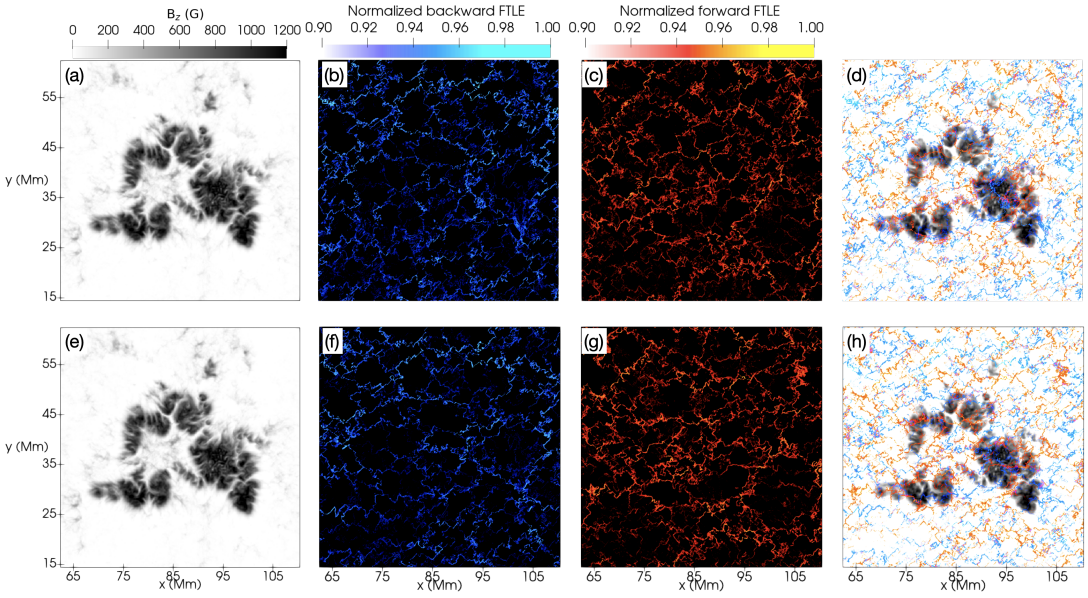

Figure 3.6: *The same as fig. 3.3, however highlighting the region from the MURaM simulation shown in fig. 2.2.*

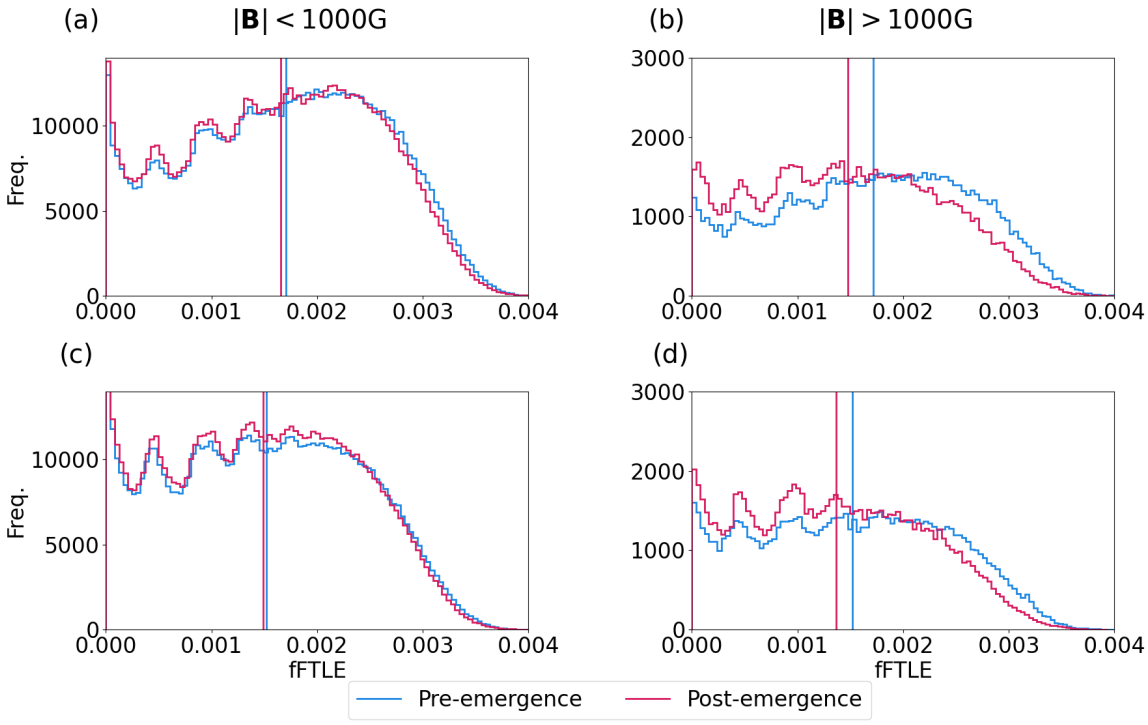

Figure 3.7: *Distribution of the 20-minute forward-FTLE field at times of 2h (prior to the emergence of the AR; blue) and 83h (post-emergence of the AR; magenta) in the MURaM simulation. The top row shows the FTLE distributions calculated using velocities from the simulation and the bottom rows shows the distributions using velocities recovered by* $\mathrm{DV}_{\mathrm{R2D2}}$*. The left column shows the distributions over a region, which contains* $<$ kG *magnetic flux density throughout the simulation and the right column shows the distributions over a region where* $>$ kG *magnetic flux density is present when the AR emerges. The vertical lines highlight the respective positive forward-FTLE distribution means.*

Chapter 4

Identifying and Tracking the Presence of Active Regions with Coherent Photospheric Flow Structures

4.1 Introduction

Space weather (SW), as mentioned in chapter 1, is of great importance. The consequences of its effects are far reaching and threaten electrical systems whose reliability is essential for maintaining modern human activity (Schwenn 2006). SW is driven by the release of magnetic flux in the solar atmosphere, nested in ARs (van Driel-Gesztelyi & Green 2015) can manifest as the appearance of dark spots in the solar photosphere called sunspots or pores, depending on their size. Forecasting SW presents many challenges (e.g., see Singer et al. 2001, Vassiliadis et al. 2007). One of these challenges is the timeliness of accurate forecasting, and an aspect of being able to do this is taking local measurements prior to an event occurring, the difficulty in this task is that it is hard to identify and classify regions where a solar flare (SF) or coronal mass ejection (CME) etc. may take place if they are not already being directly observed.

As such, many local forecasting techniques, which predict for the time of observation, are called nowcasting (Camporeale 2019). Thus it is important that techniques focus more on the forecasting aspect, and one way to achieve this may be to identify where ARs will emerge, so that they may be monitored long before the onset of any SW events.

It has been revealed through helioseismology that there are statistically significant subsurface signals prior to the emergence of ARs, namely subsurface flows in the Sun are dominated by convective motions of plasma, and that the emergence of magnetic flux occurs in the intergranular regions between supergranular cells (Birch et al. 2013). In addition to this, diverging horizontal flows have been detected up to 100min prior to emergence in observation (Toriumi et al. 2012) and up to 3h prior in simulations (Rempel & Cheung 2014). In other words, there is evidence that plasma motions at the surface level exhibit behaviours that are a direct result of the interaction between plasma and magnetic fields. As highlighted in chapter 3, the FTLE field provides an insight to changes in the flow topology in the presence of magnetic fields. Recently, it has also been evidenced by Silva et al. (2023) that the FTLE field presents signals that are indicative of the emergence of magnetic flux. This signal is shown in fig. 4.1. This novel approach determines the presence of a signal in the FTLE field up to 10h prior to the time of emergence, for example see fig. 4.2 and fig. 4.3. However, a problem with this methodology is that the identification of the FTLE field demands knowledge of the horizontal photospheric velocities. DV, as shown in chapters 2 and 3 provides the potential of recovering flows from observations and preserving structure in the FTLE field. Therefore, in this chapter, the ability for the FTLE field to indicate the presence of intense regions of magnetic flux density and track these is presented and we also determine the effectiveness of DeepVel at supporting such a method. We present that changes in the FTLE field over a field of view are directly correlated with the presence of photospheric magnetic fields, and that emulated flows from observations using DV can be used to track the location of ARs at the mesogranular scale.

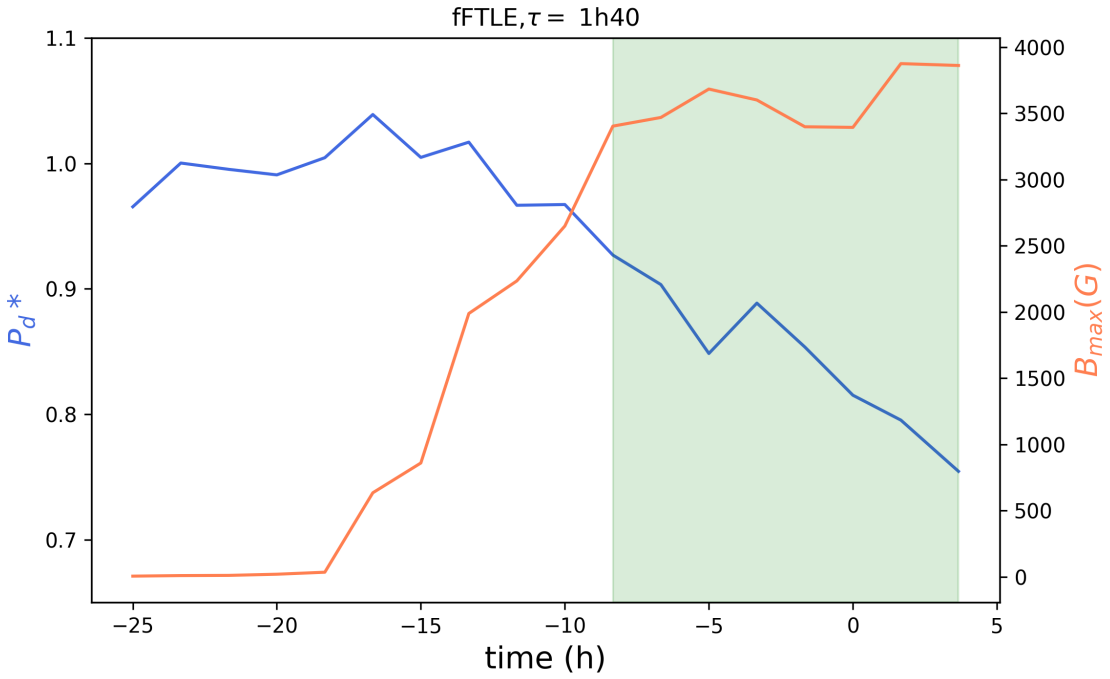

Figure 4.1: *Plot from Silva et al. (2023), showing the normalised distance between peaks in the 100-minute forward FTLE field in the R2D2 simulation around the time of emergence, which was chosen to be t_0 by the authors. The blue line corresponds to the average distance between peaks (P_d^*) of the forward FTLE ridges and the orange line shows the evolution of the maximum unsigned flux ($B_{\max}$) in Gauss. The green shaded region shows when a significant enough change in the distances between peaks to identify the presence of an emerging AR*

. Up to 6 hours prior to the emergence time of the AR, there is a distinct drop in the FTLE ridge distance.

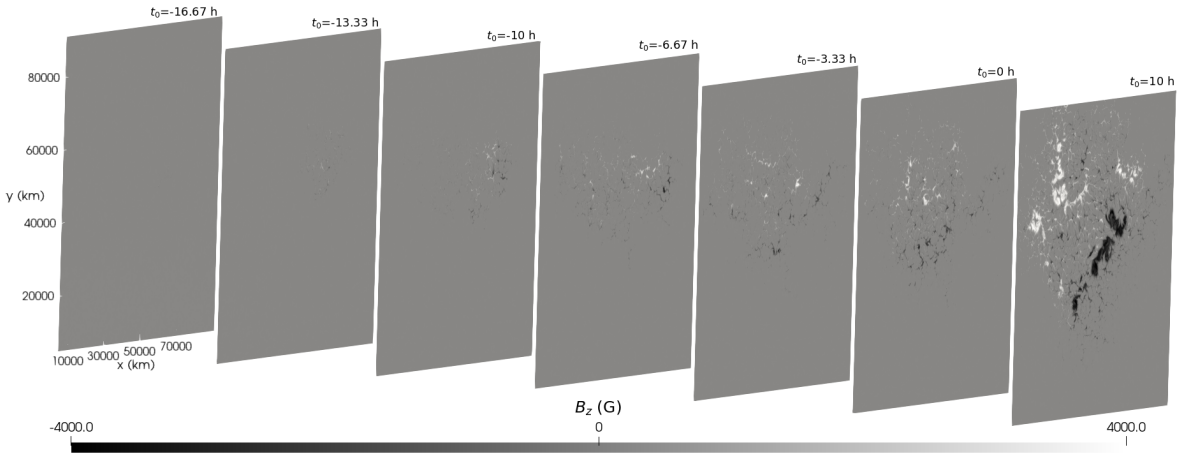

Figure 4.2: *Plot from Silva et al. (2023) showing the time-evolution of the magnetic field at the $\tau = 1$ surface from the R2D2 simulation about the time of AR emergence in the photosphere, t_0.*

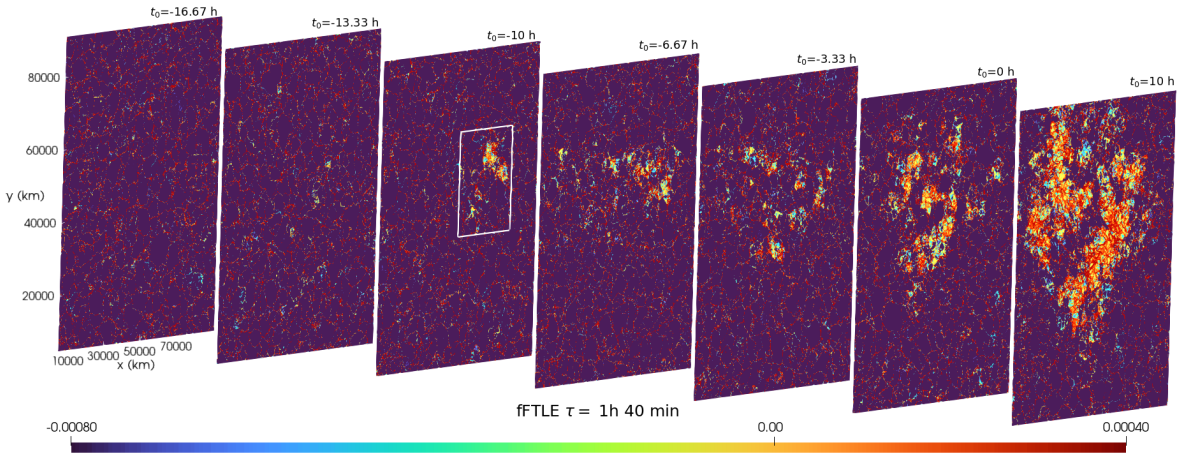

Figure 4.3: *Plot from Silva et al. (2023) showing the time-evolution of the 100-minute forward FTLE field over the photosphere about the time of AR emergence, t_0. Distinct material lines form over the region where the magnetic flux is strongest up to 10 hours before the full emergence.*

4.2 Methodology

4.2.1 The Finite-Time Lyapunov Exponent

As in chapter 3, flows from the R2D2 simulation were used to calculate the 20-minute, forward-FTLE fields over an AR, revealing structures at the granular scale. The forward-FTLE field describes diverging flow structures, which are assumed to correspond with the flows that are present during flux emergence as magnetic fields rise and cause granules to stretch and fragment. Whilst the backward-FTLE field also shows changes in the flow, these describe changes to converging flow structures which are expected to be outnumbered and dominated by the increasing strength and number of diverging features, and are therefore of less interest. This length of integration for the FTLE field was chosen mainly for two reasons. In order to maintain timeliness: the FTLE computation is expensive and can take a long time over large spatial domains, so keeping the integration over a handful of frames keeps the computational cost low as well as keeping the number of observations required small. The second reason for this is that for precise tracking: we want to identify fine-scale changes. In order to produce a useful magnetic flux detection algorithm using LCSs, the velocities it uses must be attainable at scales which discern fine structure which correlates strongly to the flux. Chapter 3 has highlighted that $\mathrm{DV}_{\mathrm{R2D2}}$ is able to reproduce the FTLE field with confidence, and detect changes in the LCSs at the granular scale.

Early in the simulation where the magnetic flux is measured to be < 1G on average, the FTLE field is representative of flow structures which correspond to the QS. From this, a mean of the positive forward-FTLE value is calculated to represent the QS. It is expected that flow structures will change locally around sites of intensifying magnetic flux as the AR emerges, thus the FTLE field will change in time. To this end, the difference in the positive forward-FTLEs, relative to the QS mean positive forward-FTLE, is calculated locally over a range of scales from the mesogranular to supergranular by rebinning the surface. Various thresholds between a relative difference of 0.05 to 1 were chosen to identify the sensitivity of the method that would produce the most accurate classification of ARs.

The version of DV used in this chapter is once again the same as that of chapter 2.

4.2.2 Identifying Magnetic structures

In order to identify regions of flux, the photospheric surface of the R2D2 simulation is rebinned to different scales, breaking the domain into tiles of fixed sizes. These are $4.8 \times 4.8\text{Mm}^2$, $9.6 \times 9.6\text{Mm}^2$, $19.2 \times 19.2\text{Mm}^2$ and $24 \times 24\text{Mm}^2$, in other words from mesogranular to supergranular scales. Individual tiles are considered to be active regions if the average unsigned flux over the region, within the tile, is calculated to be > 1kG, as per the definition from van Driel-Gesztelyi & Green (2015).

The various thresholds in the mean FTLE differences are then used to determine the presence of ARs, where if the threshold is reached, then the region is detected as active and measured against the actual magnetic activity present in the simulation.

In this dataset, only around 10–20% of the data falls in the class of AR since most of the space is dominated by granulation and prior to AR emergence, the data only contains QS-like flows. The consequence of this disparity between class sizes introduces a problem with bias—if the classifier performs well at only identifying true negatives, then the measured performance is skewed and will make the method appear to work well. However, in the method presented here, the classification is based on a simple threshold, so the classifier itself cannot be influenced by the data like typical ML-based classifiers which can be prone to overfitting. In order to correctly analyse the performance of the classifier, a number of balanced metrics should be considered which are not skewed by class size, described in the next section.

4.3 Results

Here, the results of using a threshold on the difference of the mean positive forward-FTLE over subregions of a simulated active region are presented. In order to determine the success of the FTLE field as a classifier of the presence of intense magnetic fluxes, confusion matrices have been calculated for various times and for the various thresholds.

Confusion matrices display the correctly and incorrectly classified cases, these are true positives (TP: the correctly classified positive cases), true negatives (TN: the correctly classified negative cases), false positives (FP: incorrectly classified negative cases) and false negatives (FN: the incorrectly identified positive cases). A positive classification in this case is one in which a region is considered active. From the various types of the classifications, the accuracy, precision, recall and F1-score are all calculated. These metrics are defined and interpreted as follows:

- Accuracy: $\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FN}+\mathrm{FP}}$, the ratio of correctly classified cases, i.e. the probability of correctly classifying a region as active or quiet.
- Balanced Accuracy: $\frac{1}{2}\left(\frac{TP}{TP+FN}+\frac{TN}{TN+FP}\right)$, the average of the true positive and true negative rates (see below). This gives a more balanced value when considering cases where the number of cases in one of the classes is imbalanced.
- Precision: $\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}$, the proportion of TPs of the predicted positive cases, i.e. this is the probability of the case being positive given that it has been identified as positive.
- Recall (or TPR): $\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$, the ratio of positive cases identified of all positive cases, or the probability of the classifier giving a positive result, given that the case is in fact positive.
- FPR: $\frac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TN}}$, the ratio of negative cases identified of all negative cases, or the probability of the classifier giving a negative result, given that the case is in fact negative.
- F1-score: $2\times\frac{\mathrm{precision}\times\mathrm{recall}}{\mathrm{precision}+\mathrm{recall}}$, the harmonic mean of the precision and recall scores,

where the TPR and FPR are the true positive rate and the false positive rate, respectively. The balanced accuracy, precision, recall and F1-score are not influenced by the class imbalance as they take into account the class size in their definitions.

For the best case (considered to be the highest F1-score), the TPRs and FPRs were plotted on the receiver operator characteristic curves shown in fig. 4.6.

Table 4.1: *Performance of various relative FTLE mean differences, calculated from the R2D2 simulation (Hotta & Iijima 2020) velocities over various grid spacings as a classifier for AR levels of magnetic flux.*

Scale (Mm)	FTLE (Relative difference)	Metric				
		Accuracy	Balanced Accuracy	Precision	Recall	f1 Score
	0.00	0.118	0.500	0.118	1.000	0.210
	0.05	0.425	0.656	0.165	0.957	0.281
	0.10	0.648	0.773	0.242	0.936	0.384
	0.15	0.800	0.822	0.354	0.851	0.500
	0.20	0.873	0.808	0.472	0.723	0.571
	0.25	0.920	0.807	0.660	0.660	0.660
	0.30	0.923	0.762	0.722	0.553	0.627
	0.35	0.928	0.738	0.821	0.489	0.613
	0.40	0.935	0.733	0.957	0.468	0.629
	0.45	0.935	0.723	1.000	0.447	0.618
4.8	0.50	0.930	0.702	1.000	0.404	0.576
	0.55	0.913	0.628	1.000	0.255	0.407
	0.60	0.903	0.585	1.000	0.170	0.291
	0.65	0.893	0.543	1.000	0.085	0.157
	0.70	0.890	0.532	1.000	0.064	0.120
	0.75	0.888	0.521	1.000	0.043	0.082
	0.80	0.883	0.500	0.000	0.000	0.000
	0.85	0.883	0.500	0.000	0.000	0.000
	0.90	0.883	0.500	0.000	0.000	0.000
	0.95	0.883	0.500	0.000	0.000	0.000

Continued on next page

Scale (Mm)	FTLE (Relative difference)	Metric				
		Accuracy	Balanced Accuracy	Precision	Recall	f1 Score
	1.00	0.883	0.500	0.000	0.000	0.000

Continued on next page

Scale (Mm)	FTLE (Relative difference)	Metric				
		Accuracy	Balanced Accuracy	Precision	Recall	f1 Score
	0.00	0.130	0.500	0.130	1.000	0.230
	0.05	0.420	0.667	0.183	1.000	0.310
	0.10	0.670	0.745	0.262	0.846	0.400
	0.15	0.830	0.837	0.423	0.846	0.564
	0.20	0.850	0.718	0.438	0.538	0.483
	0.25	0.890	0.740	0.583	0.538	0.560
	0.30	0.900	0.714	0.667	0.462	0.545
	0.35	0.890	0.642	0.667	0.308	0.421
	0.40	0.910	0.654	1.000	0.308	0.471
	0.45	0.910	0.654	1.000	0.308	0.471
9.6	0.50	0.900	0.615	1.000	0.231	0.375
	0.55	0.900	0.615	1.000	0.231	0.375
	0.60	0.890	0.577	1.000	0.154	0.267
	0.65	0.880	0.538	1.000	0.077	0.143
	0.70	0.870	0.500	0.000	0.000	0.000
	0.75	0.870	0.500	0.000	0.000	0.000
	0.80	0.870	0.500	0.000	0.000	0.000
	0.85	0.870	0.500	0.000	0.000	0.000
	0.90	0.870	0.500	0.000	0.000	0.000
	0.95	0.870	0.500	0.000	0.000	0.000
	1.00	0.870	0.500	0.000	0.000	0.000

Continued on next page

Scale (Mm)	FTLE (Relative difference)	Metric				
		Accuracy	Balanced Accuracy	Precision	Recall	f1 Score
	0.00	0.080	0.500	0.080	1.000	0.148
	0.05	0.360	0.652	0.111	1.000	0.200
	0.10	0.560	0.761	0.154	1.000	0.267
	0.15	0.760	0.870	0.250	1.000	0.400
	0.20	0.840	0.685	0.250	0.500	0.333
	0.25	0.880	0.707	0.333	0.500	0.400
	0.30	0.920	0.728	0.500	0.500	0.500
	0.35	0.920	0.728	0.500	0.500	0.500
	0.40	0.880	0.478	0.000	0.000	0.000
	0.45	0.880	0.478	0.000	0.000	0.000
19.2	0.50	0.920	0.500	0.000	0.000	0.000
	0.55	0.920	0.500	0.000	0.000	0.000
	0.60	0.920	0.500	0.000	0.000	0.000
	0.65	0.920	0.500	0.000	0.000	0.000
	0.70	0.920	0.500	0.000	0.000	0.000
	0.75	0.920	0.500	0.000	0.000	0.000
	0.80	0.920	0.500	0.000	0.000	0.000
	0.85	0.920	0.500	0.000	0.000	0.000
	0.90	0.920	0.500	0.000	0.000	0.000
	0.95	0.920	0.500	0.000	0.000	0.000
	1.00	0.920	0.500	0.000	0.000	0.000

Continued on next page

Scale (Mm)	FTLE (Relative difference)	Metric				
		Accuracy	Balanced Accuracy	Precision	Recall	f1 Score
	0.00	0.125	0.500	0.125	1.000	0.222
	0.05	0.438	0.679	0.182	1.000	0.308
	0.10	0.625	0.786	0.250	1.000	0.400
	0.15	0.688	0.607	0.200	0.500	0.286
	0.20	0.875	0.714	0.500	0.500	0.500
	0.25	0.875	0.714	0.500	0.500	0.500
	0.30	0.875	0.714	0.500	0.500	0.500
	0.35	0.875	0.714	0.500	0.500	0.500
	0.40	0.875	0.500	0.000	0.000	0.000
	0.45	0.875	0.500	0.000	0.000	0.000
24.0	0.50	0.875	0.500	0.000	0.000	0.000
	0.55	0.875	0.500	0.000	0.000	0.000
	0.60	0.875	0.500	0.000	0.000	0.000
	0.65	0.875	0.500	0.000	0.000	0.000
	0.70	0.875	0.500	0.000	0.000	0.000
	0.75	0.875	0.500	0.000	0.000	0.000
	0.80	0.875	0.500	0.000	0.000	0.000
	0.85	0.875	0.500	0.000	0.000	0.000
	0.90	0.875	0.500	0.000	0.000	0.000
	0.95	0.875	0.500	0.000	0.000	0.000
	1.00	0.875	0.500	0.000	0.000	0.000

Table 4.2: *Performance of various relative FTLE mean differences, calculated using velocities recovered by DV from the R2D2 simulation (Hotta & Iijima 2020) intensities, over various grid spacings as a classifier for AR levels of magnetic flux.*

Scale (Mm)	FTLE (Relative difference)	Metric				
		Accuracy	Balanced Accuracy	Precision	Recall	f1 Score
4.8	0.00	0.118	0.500	0.118	1.000	0.210
	0.05	0.475	0.684	0.178	0.957	0.300
	0.10	0.723	0.806	0.287	0.915	0.437
	0.15	0.853	0.852	0.435	0.851	0.576
	0.20	0.893	0.828	0.530	0.745	0.619
	0.25	0.910	0.801	0.608	0.660	0.633
	0.30	0.923	0.762	0.722	0.553	0.627
	0.35	0.940	0.763	0.926	0.532	0.676
	0.40	0.933	0.731	0.917	0.468	0.620
	0.45	0.935	0.723	1.000	0.447	0.618
	0.50	0.928	0.691	1.000	0.383	0.554
	0.55	0.915	0.638	1.000	0.277	0.433
	0.60	0.905	0.596	1.000	0.191	0.321
	0.65	0.898	0.564	1.000	0.128	0.226
	0.70	0.888	0.521	1.000	0.043	0.082
	0.75	0.885	0.511	1.000	0.021	0.042
	0.80	0.885	0.511	1.000	0.021	0.042
	0.85	0.883	0.500	0.000	0.000	0.000
	0.90	0.883	0.500	0.000	0.000	0.000
	0.95	0.883	0.500	0.000	0.000	0.000

Continued on next page

Scale (Mm)	FTLE (% difference)	Metric				
		Accuracy	Balanced Accuracy	Precision	Recall	f1 Score
	1.00	0.883	0.500	0.000	0.000	0.000

Continued on next page

Scale (Mm)	FTLE (% difference)	Metric				
		Accuracy	Balanced Accuracy	Precision	Recall	f1 Score
9.6	0.00	0.130	0.500	0.130	1.000	0.230
	0.05	0.470	0.630	0.177	0.846	0.293
	0.10	0.750	0.791	0.324	0.846	0.468
	0.15	0.830	0.771	0.409	0.692	0.514
	0.20	0.850	0.718	0.438	0.538	0.483
	0.25	0.890	0.740	0.583	0.538	0.560
	0.30	0.900	0.714	0.667	0.462	0.545
	0.35	0.920	0.692	1.000	0.385	0.556
	0.40	0.910	0.654	1.000	0.308	0.471
	0.45	0.900	0.615	1.000	0.231	0.375
	0.50	0.900	0.615	1.000	0.231	0.375
	0.55	0.890	0.577	1.000	0.154	0.267
	0.60	0.880	0.538	1.000	0.077	0.143
	0.65	0.880	0.538	1.000	0.077	0.143
	0.70	0.870	0.500	0.000	0.000	0.000
	0.75	0.870	0.500	0.000	0.000	0.000
	0.80	0.870	0.500	0.000	0.000	0.000
	0.85	0.870	0.500	0.000	0.000	0.000
	0.90	0.870	0.500	0.000	0.000	0.000
	0.95	0.870	0.500	0.000	0.000	0.000
	1.00	0.870	0.500	0.000	0.000	0.000

Continued on next page

Scale (Mm)	FTLE (% difference)	Metric				
		Accuracy	Balanced Accuracy	Precision	Recall	f1 Score
19.2	0.00	0.080	0.500	0.080	1.000	0.148
	0.05	0.520	0.739	0.143	1.000	0.250
	0.10	0.680	0.826	0.200	1.000	0.333
	0.15	0.800	0.891	0.286	1.000	0.444
	0.20	0.880	0.935	0.400	1.000	0.571
	0.25	0.880	0.707	0.333	0.500	0.400
	0.30	0.920	0.728	0.500	0.500	0.500
	0.35	0.880	0.478	0.000	0.000	0.000
	0.40	0.880	0.478	0.000	0.000	0.000
	0.45	0.880	0.478	0.000	0.000	0.000
	0.50	0.920	0.500	0.000	0.000	0.000
	0.55	0.920	0.500	0.000	0.000	0.000
	0.60	0.920	0.500	0.000	0.000	0.000
	0.65	0.920	0.500	0.000	0.000	0.000
	0.70	0.920	0.500	0.000	0.000	0.000
	0.75	0.920	0.500	0.000	0.000	0.000
	0.80	0.920	0.500	0.000	0.000	0.000
	0.85	0.920	0.500	0.000	0.000	0.000
	0.90	0.920	0.500	0.000	0.000	0.000
	0.95	0.920	0.500	0.000	0.000	0.000
	1.00	0.920	0.500	0.000	0.000	0.000

Continued on next page

Scale (Mm)	FTLE (% difference)	Metric				
		Accuracy	Balanced Accuracy	Precision	Recall	f1 Score
24.0	0.00	0.125	0.500	0.125	1.000	0.222
	0.05	0.563	0.750	0.222	1.000	0.364
	0.10	0.688	0.821	0.286	1.000	0.444
	0.15	0.750	0.643	0.250	0.500	0.333
	0.20	0.875	0.714	0.500	0.500	0.500
	0.25	0.875	0.714	0.500	0.500	0.500
	0.30	0.875	0.714	0.500	0.500	0.500
	0.35	0.813	0.464	0.000	0.000	0.000
	0.40	0.813	0.464	0.000	0.000	0.000
	0.45	0.875	0.500	0.000	0.000	0.000
	0.50	0.875	0.500	0.000	0.000	0.000
	0.55	0.875	0.500	0.000	0.000	0.000
	0.60	0.875	0.500	0.000	0.000	0.000
	0.65	0.875	0.500	0.000	0.000	0.000
	0.70	0.875	0.500	0.000	0.000	0.000
	0.75	0.875	0.500	0.000	0.000	0.000
	0.80	0.875	0.500	0.000	0.000	0.000
	0.85	0.875	0.500	0.000	0.000	0.000
	0.90	0.875	0.500	0.000	0.000	0.000
	0.95	0.875	0.500	0.000	0.000	0.000
	1.00	0.875	0.500	0.000	0.000	0.000

The metrics described in the beginning of this section were calculated for a range of

different spatial scales and thresholds of the relative forward-FTLE mean difference. The performance metrics for the classifier when using simulation velocities is presented in table 4.1, and those using the velocities predicted by DV_{R2D2} are shown in table 4.2. Both tables show that the 20-minute forward-FTLE fields perform the best as a classifier for the presence of ARs at mesogranular scales of $4.8 \times 4.8 Mm^2$, where the scores for all performance metrics achieve their highest values. These forward-FTLE fields correspond to diverging features on the granular scale, which will have a response to the presence of emerging magnetic fields during AR formation. Thus, the response of granular features will correspond to changes in the magnetic field at some scale. Given that, during emergence, the magnetic field does not dominate the entire region, larger regions of sparse magnetic field will be required to amass a notable change in flow structures. This result suggests flows on the granular scale have a persisting influence on magnetic features on the mesogranular scale during AR formation or vice versa.

The simulation velocities support that a relative threshold in the range of 0.25–0.35 provide the best balanced accuracies. In fact the best is the threshold of 0.25 for maintaining a strong f1-score. This value has been selected as the strongest since we wish to optimise the recall, while still maintaining a good precision. In fact, in this case, the recall is considered to be more important than the precision, since the target is to identify the location of an AR, and the cost of mislabelling a local quiet region is less than that of mislabelling a local AR. Hence, the balanced precision and recall scores of the relative threshold of 0.25 supports the best overall detection.

Interestingly, DV_{R2D2} provides a stronger performance measured by the balanced accuracy up until a relative threshold of 0.25. Additionally, the precision and recall, although similar, are also greater for most of the thresholds using DV velocities. The changes in these values is typical as the increasing threshold will cause less positives to be identified, but when they are they are more likely to be true positives. Thus, as the threshold increases, the TPR (recall) will decrease since it is less likely for a region to be classified as positive (active) and the precision therefore increases as the likelihood of detecting a true positive increases.

The precision and recall scores are roughly balanced for both at a threshold value 0.25. The improved performance from DV over the simulation velocities is somewhat unexpected, however the FTLE fields produced from DV appear to show a slightly reduced difference in quiet and active region means, compared those produced by simulation velocities (see fig. 3.4). Whilst this is a small error, this difference in sensitivity is likely enough to to change how likely it is for a region to be classified as containing an AR when using the DV velocities.

Other scales in the tables become a bit more difficult to interpret due to the fact that for patches of $19.2 \times 19.2\text{Mm}^2$ and $24 \times 24\text{Mm}^2$ there are only 25 and 16 tiles over the whole FOV and only a few (2-5) patches considered active in terms of the levels of magnetic flux at these scales. Thus, identifying larger regions amplifies the significance of any mislabelling of patches. Additionally, since changes in the forward-FTLE value represents features on the granular scale, any changes due to the presence of magnetic flux are expected to be highly localised and hence the chance of error in classifying a region correctly is expected to increase.

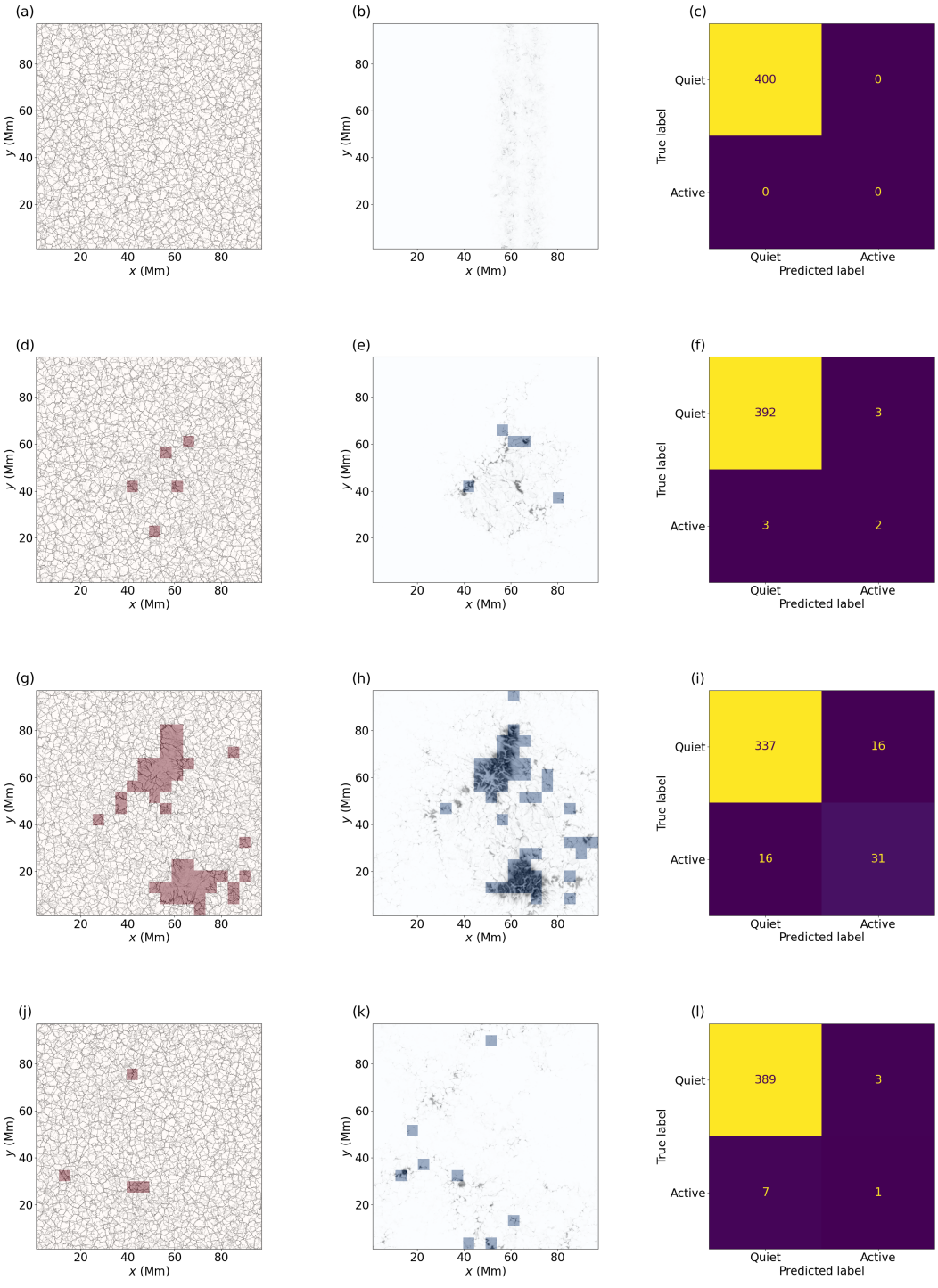
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
y (Mm)
x (Mm)
20
40
60
80
Quiet
Active
True label
Predicted label
400
0
0
0
392
3
3
2
337
16
16
31
389
3
7
1

Figure 4.4 ***(previous page)*****:** *The 20-minute forward-FTLE fields from R2D2 simulation velocities (left column), magnetic flux (middle) and confusion matrices of the regions classified as active via the FTLE field vs the magnetic field (right). The top, second, third and fourth rows display key times throughout the sunspot evolution of 3 (mostly quiet), 40 (emerging sunspot), 60 (peak photospheric magnetic flux) and 170 (near full decay of sunspot) hours. Cases identified as active via the FTLE field are highlighted in pink when the FTLE mean value differs by 0.25 from the mean FTLE value in the QS. The actual positive cases, determined by the concentration of magnetic flux, are highlighted in blue.*

Figure 4.4 and fig. 4.5 display examples of the AR classifier for simulation FTLEs and $\mathrm{DV}_{\mathrm{R2D2}}$, respectively. It is clear that in both instances, where little magnetic flux is present (see panels a, b c, d, e, f, j, k and l), the FTLE fields highlight changes in regions close to those where $>$ kG concentrations of magnetic flux are present. Despite their location not being exact, roughly the same number of ARs are detected by the FTLE field as their are measured by the averaged magnetic flux density, early in emergence phase. When the presence of the magnetic flux is at its peak, changes in the FTLE field coincide largely with the ARs. Namely, the true positive rate hits 0.67. So for large ARs, 2/3s of the entire region will be detected correctly. In addition to this, both the R2D2 and $\mathrm{DV}_{\mathrm{R2D2}}$ flows detect a similar number of positives as measured by the levels of flux. It is apparent that these regions, where they do not coincide exactly with the presence of the AR, are very close to the ARs.

The above highlights two major points. Firstly, changes in the flow topology, determined by the FTLE field, are indicative of the presence of magnetic flux and can confidently locate ARs on the mesogranular scale. Secondly, the NN DV has the capacity to predict accurate enough flows, at least from simulation, to provide almost an excellent performance in recovering flow topology, closely matching that identified within simulations.

The receiver operating characteristic (ROC; which refers to the performance of a detector) curve shows the trade off between sensitivity (measured by the TPR) and the FPR (1 - specificity). The area under the curve (AUC) presents a measure for the performance as a value between 0 and 1. When the AUC is close to 1, the classifier is correctly identifying true positives almost all the time, if the AUC is close to 0.5 (shown by the dashed line in fig. 4.6) then the classifier is randomly guessing the labels. In the best case scenario, the sensitivity is

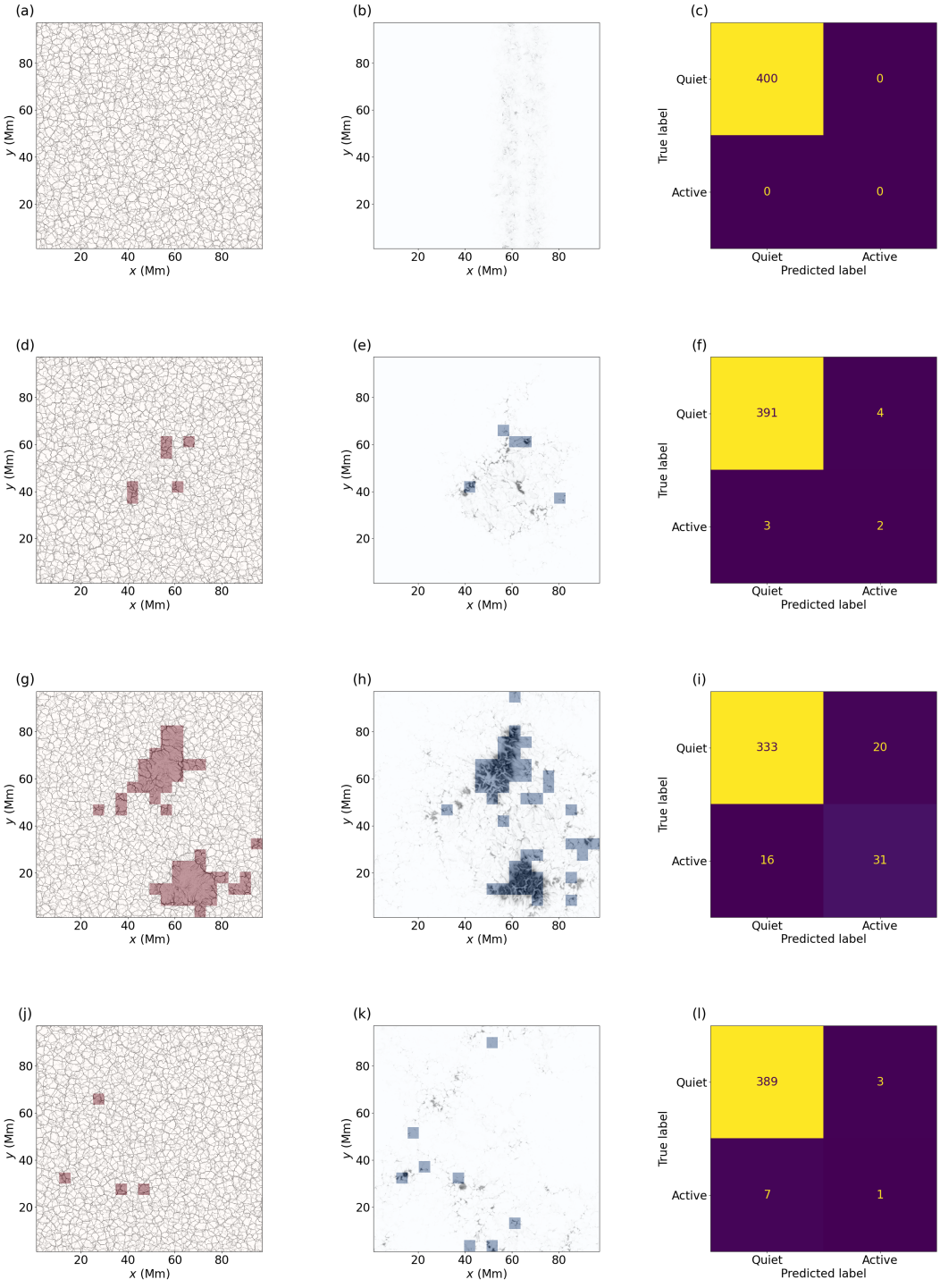

Figure 4.5: *The same as fig. 4.4, however the velocities used for calculating the FTLEs were estimated by DV.*

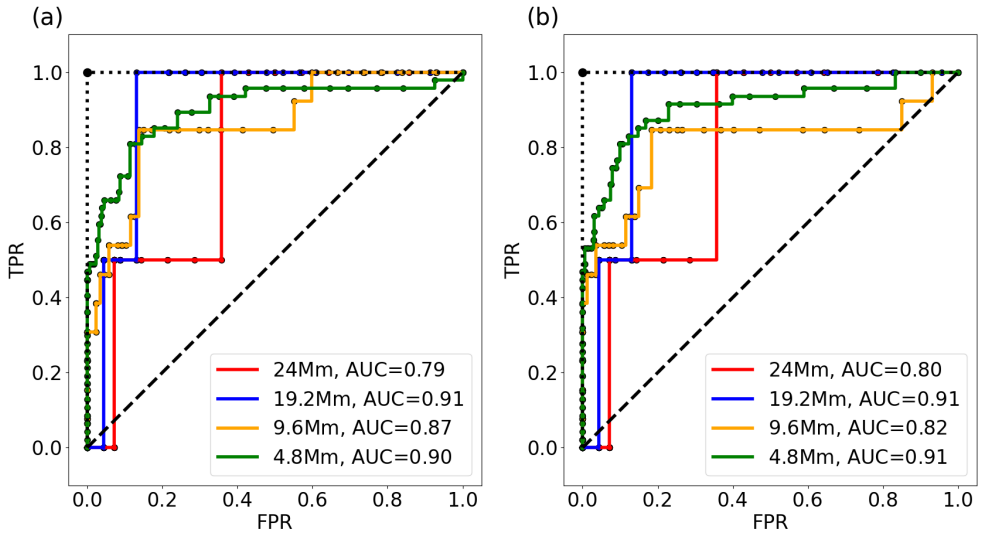

Figure 4.6: *ROC curves for 20 evenly spaced thresholds of the relative positive forward-FTLE mean difference and 4 different spatial scales (depicting the length of one side of the tile). Each point along the curve shows the FPR and TPR for each threshold. Each curve is labelled with the area under the curve (AUC). The diagonal, dashed line represents the possible case and the dotted line shows the best case scenario, where the optimum case is highlighted by the black point. Panel (a) shows the curves when using the velocities from the R2D2 simulation and (b) shows the curves when using velocities estimated by DV for calculating the FTLEs.*

1 and the specificity is 1, this curve and point are also shown by the dotted line in the figure.

The performance when using both the R2D2 and the DV velocities is very similar for the FTLE mean difference threshold of 0.25 at a $4.8 \times 4.8\text{Mm}^2$ scale. This is highlighted by table 4.1 and table 4.2. We see that the ROC curves take on similar shapes with DV showing better specificity at certain thresholds. Whilst the green curves, representing the smallest studied tile size, does not outperform the larger tiles of $19.2 \times 19.2\text{Mm}^2$ as measured by the AUC, the classification has a smaller balanced accuracy and a much smaller f1 score, highlighting that the apparent success of the ROC curve is likely due to the imbalance in the classes. Thus it is decided that the $4.8 \times 4.8\text{Mm}^2$ tiling works best for classifying using granules.

4.4 Summary

In this chapter, the pipeline for new method was proposed for the detection of ARs. Using velocities from simulation, it was shown that to a high degree of accuracy the FTLE field can indicate the presence of high concentrations ($>$ kG) of magnetic flux. In addition to this, it was highlighted that the detection of magnetic flux can be revealed with only a short time series (20 minutes) of continuum images, from which the DeepVel network can be used to predict the corresponding flow fields, which can then be used to derive the FTLE field, given the measurement of some baseline FTLE value for quiet Sun flows.

An advantage to this approach is that the classifier itself is based on a threshold relating to physical properties of the flow, which are expected to correlate to magnetic structures. This means that this approach, once fine-tuned, is stable across multiple datasets and is highly reproducible and interpretable. However, it is also possible that the values that this threshold is based on change from dataset to dataset and thus may in fact fail as it isn't adaptive, therefore needing to be analysed and tuned for each new dataset.

Whilst the proposed method is not exact in its accuracy for detecting the location of flux, it is able to determine its presence almost always, and does not falsely produce positive classifications when there is no significant magnetic flux present in the photosphere of R2D2 simulation. It is possible that this classifiers performance is good because it more often than not predicts negative cases, but the balance of precision and recall at the identified best threshold inspire confidence. The heavily weighted bias in class sizes indicate that threshold is doing a significantly better job than random guessing, which would require a precision of $\approx$ 0.1–0.2. In order to fully test this method, further studies should be carried out by applying this to QS only simulations to distinguish whether or not this can be used to always detect the presence of magnetic flux without causing false alarm. Furthermore, this method, in tandem with the method presented by Silva et al. (2023) showcase a significant step toward being able to efficiently forecast and locate the presence of emerging ARs. Thus establishing a method that would provide ample time for researchers to collect high quality observations

of complex magnetic structures.

Chapter 5

Identifying Coherent Swirling Structures in the Chromosphere

5.1 Introduction

A major part of understanding the solar atmospheric structure and dynamics is understanding how energy is transported between layers. Each layer of the solar atmosphere has a strong influence on the next.

The chromosphere is a region of particular interest as it is home to some highly complex structures thanks to the lower chromosphere being dominated by gas pressure ($\beta > 1$) and the upper chromosphere being magnetically dominated ($\beta < 1$), see fig. 1.7. Below the chromosphere is the photosphere, in which plasma motions largely influence the structures present. At the top of the chromosphere is the transition region into the corona, where magnetic fields constantly change and influence the flow of plasma into the solar wind. In the transition region, there are steep gradients in both temperature and plasma density.

One particular area of interest in the literature is the coronal heating problem (see Klimchuk 2006), which highlights the steep gradients in atmospheric properties, as the temperature rapidly increases, whilst the density drops. This raises the question: what mechanisms are responsible for heating the corona? One supposed mechanism for the heating of the corona

is the influence of vortical motions. For example, photospheric vortices have been linked to the coronal heating problem (e.g., see Zirker 1993). Examples of chromospheric vortices and corresponding photospheric features (magnetic bright points) are highlighted in fig. 5.1. Vortices and other swirling motions have been linked to the transport of energy by building up tension in the magnetic field through the twisting of magnetic flux tubes (e.g., see Tziotziou et al. 2023). For example, in the aforementioned figure, swirling motions in the chromosphere surround regions of concentrated magnetic flux in the photosphere (Shelyag, S. et al. 2011), which are suggested to be magnetic foot points for flux tubes. Figure 5.1 highlights the Doppler shifts upflows in the vortex centres and downflows surrounding this.

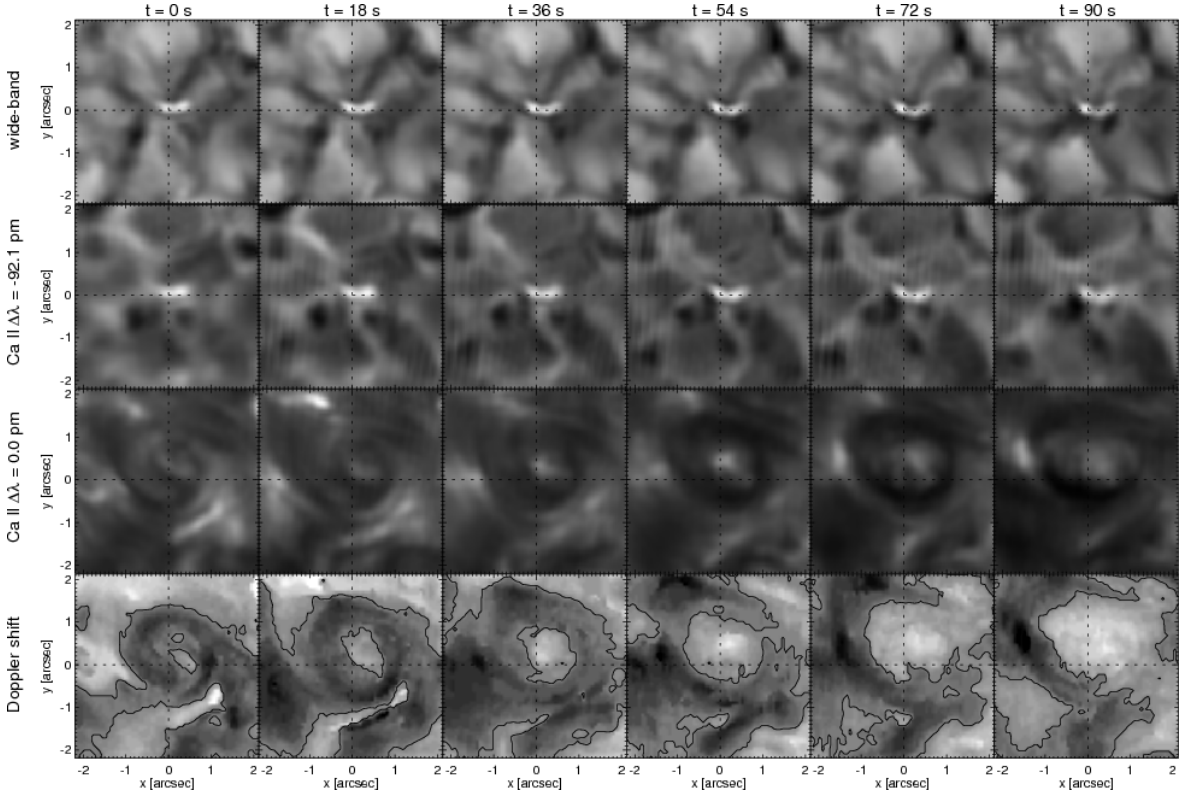

Figure 5.1: *Image taken from Wedemeyer-Böhm, S. & Rouppe van der Voort, L. (2009) showing the swirling motions in the chromosphere with corresponding photospheric bright points. The top 3 rows show images taken in the SST CRISP wide-band (photosphere), Ca II IRT 8452Å line wing (photosphere), and line core (chromosphere) respectively. The bottom row shows Doppler measurements.*

Vortices, whilst easy to identify visually, are tricky to define (see Haller et al. 2016). A simple definition may be established by finding regions of non-zero vorticity in a velocity field,

however this approach is flawed as shear flows exhibit a non-zero vorticity. A strong definition for coherent vortices will account for material invariance, that is the vortex will persist and remain detectable regardless of the frame of reference. For example, in a rotating frame of reference, defined in an Eulerian framework, a vortex may appear still. One definition from Lugt (1979) is that a vortex is 'a multitude of material particles rotating around a common center'. This definition touches on the idea that a vortex should be defined by particle trajectories and not just by the velocity field. A number of vortex detection methods exist for identifying coherent vortices in flows and these take either an Eulerian or a Lagrangian approach. These two approaches are distinct in their methodologies. Eulerian methods detect vorticity in the instantaneous velocity field with an additional quantity, derived from the velocity gradient tensor, that is used to distinguish whether the vorticity arises from a shearing motion or genuine rotations. Some of these Eulerian methods/quantities are highlighted below. The Lagrangian approach takes a more objective and frame-independent approach by tracking particle motions that result from an integration over a time-dependent velocity field. The particle trajectories are used to define contours, which can then identify the strength and locations of coherent vortical structures.

Two common Eulerian vortex detection methods are the λ_2 (see Jeong & Hussain 1995) and Γ (see Graftieaux et al. 2001) criterion. The first is able to detect vortex cores only whilst the latter is able to detect vortex boundaries also. Another method is the swirling strength criterion which identifies rotational motions as the imaginary component of the eigenvalues of the gradient tensor (see Zhou et al. 1999). Objective vortex detection requires a Lagrangian formation such as the Lagrangian averaged vorticity deviation (LAVD; see Haller et al. 2016). A common theme with vortex detection methods is that they require knowledge of the velocity field. An exception to this are methods which determine the presence of vortices by identifying morphological characteristics in highly contrasting images (e.g., see Dakanalis et al. 2021)

Studying structures within the solar atmosphere is a difficult task due to several factors. The first being that the chromosphere is rather elusive and visible only in a sparse number of wavelengths, which span a wide band and cover a wide range of heights in the chromosphere

(see fig. 1.6) , thus making it difficult to study coherent structures. The second reason is that much of the chromosphere is dominated by magnetic fields, thus large scale changes occur rapidly, making features difficult to track and see. However, vortices tend to be optically thick and therefore dense—this indicates that motions within vortical structures are easier to identify than in the background of the chromosphere (see examples from Tziotziou et al. 2023). Thirdly, the chromosphere also transitions from local thermal equilibrium (LTE) to non-LTE processes. That is, diffusive processes become important in the transfer of energy in the chromosphere and thus images are much less coherent. This also creates difficulty in modelling chromospheric processes. In recent times a number of numerical codes for modelling the chromosphere have appeared, for example Bifrost (see Gudiksen et al. 2011), MANCHA (see Khomenko & Collados 2006) and most recently a chromospheric extension of the MURaM code (see Przybylski et al. 2022). Retrieving plasma motions from the chromosphere is difficult for the reasons mentioned above, as standard optical tracking techniques and others cannot be applied. The ability for DV for recovering flows at various different levels of the solar atmosphere has already been presented for simulations of the QS and penumbral flows (see Tremblay & Attie 2020). The previous work highlighted a performance for lower resolution images and longer cadences designed to mimic the quality of observations from SDO/HMI. This, combined with the recent development of high resolution simulations and the development of numerical codes for synthesising chromospheric imagery, such as Lightweaver (a non-LTE RT code for synthesising images; see Osborne & Milić 2021), indicates that for the first time it may be possible to reproduce coherent flow structures in the chromosphere.

In this chapter, a cascade of two neural networks is proposed in which the outputs of one version of DV is used as the inputs for a second version. By using the radiation formation heights of synthesised chromospheric images from the Bifrost code, a version of DV is trained to predict velocities in chromospheric imagery where swirling structures are present. Thus what is presented is a first attempt at performing flow recovery for coherent structure identification in the chromosphere using DV.

5.2 Methods

5.2.1 Data

As in chapter 2, DV was trained to recover velocities from simulated data of the solar atmosphere as these provide access to a ground truth for velocities. The simulation in question was the run of the Bifrost (Gudiksen et al. 2011) simulation labelled 'ch024031'[1], which models the solar atmosphere within a region containing a coronal hole using LTE modelling for hydrogen in the atmosphere to improve computational efficiency (e.g., see Wedemeyer et al. 2022). As described above, the chromosphere is home to non-LTE processes, thus this simulation is not a true representation of chromospheric dynamics and attempts to emulate the chromosphere using LTE-modelling for efficiency, for proof-of-concept modelling and testing of methods it provides a good, however flawed, starting point. The simulation contains no large magnetic fluxes, however a horizontal magnetic field of 200G is steadily fed into the photospheric downflows from the bottom boundary. As the field strength slowly increases, swirling motions become apparent at chromospheric heights of around 1–2Mm. The speeds of various layers are shown in fig. 5.2, highlighting that chromospheric swirling structures may form over regions of QS granulation (i.e. only small-scale vortices are present, if any) in the lower layers of the atmosphere as suggested by the bright point theory.

Images of the chromosphere were synthesised using the lightweaver code, to match the 3mm band seen by the Atacama large millimetre array (ALMA) radio telescope, which observes the Sun in submillimetre to millimetre wavelengths. The 3mm band is in the microwave/far infrared (FIR) spectrum, and images cover a wide range of heights, seeing almost the entire depth of the chromosphere. Because of this, synthesised 3mm band images provide a reasonable testing ground for the applicability of DV in the chromosphere. It should be noted that images from ALMA in this spectrum have a relatively low range resolutions (1.4–2.1") compared to most modern telescopes like SDO which sees at 1" and below. Whereas the synthesised images from Bifrost have a pixel resolution of 31.25kmpix^{-1} (corresponding

[1]Available for download at this URL: `https://sdc.uio.no/search/simulations`

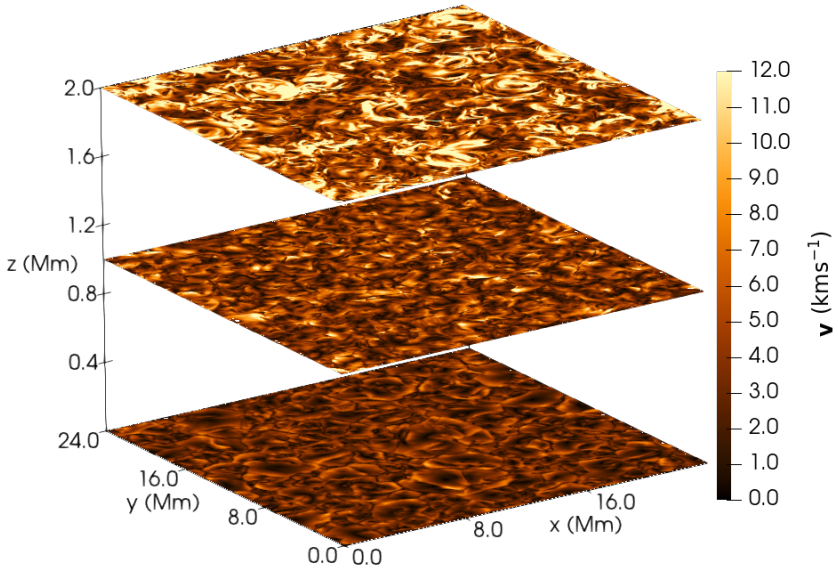

Figure 5.2: *Magnitude of the horizontal velocities from the Bifrost simulation 'ch024031'. These are shown at heights of* 0Mm, 1Mm *and* 2Mm *relative to the surface of the Sun, displaying the photosphere, mid-chromosphere and upper-chromosphere.*

to seeing with an angular resolution of around 0.08"), which is more in line with the effective resolutions of images available from SST (see Rouppe van der Voort, L. H. M. et al. 2020, for details). The images are separated with a cadence of 10s, which is on the order of magnitude of the evolution time of images on the Mm scale (Wedemeyer-Böhm et al. 2005), thus we expect that enough information is available that DV can discern coherent patterns from one image to the next. Because of this longer-than-ideal cadence, images were rebinned to half the resolution, removing the finest scale and therefore most rapidly changing structures.

As mentioned above, the synthesised 3mm images cover a great range of depths. As such, many features present in images exist at largely different vertical locations despite appearing to be close together in the horizontal plane. Due to the transient nature of much of the chromosphere, the formation heights of the 3mm band were also used as input for DV. It was suspected that the heights will provide valuable material for DV to be able to recover key structures in velocity fields. It is also known that these heights cannot be identified from

observations directly, thus for this network to be useful a way of identifying the heights needs to be determined. As such a second copy of DV was trained purely to recover heights from the 3mm band images.

5.2.2 DeepVel

Consecutive pairs of sub-images, extracted from the synthesised 3mm band images, were chosen at random from 618 available times. The sub-images chosen were of size $3.125 \times 3.125\text{Mm}^2$ ($50 \times 50\text{pix}^2$) of the $24 \times 24\text{Mm}^2$ full size images. 1500 pairs were used for training, a further 300 were used for validation and a further 300 were used for testing the network by hand. The results of the network performance for the test set are shown in the following section. Contained in these pairs were the 3mm intensities, the formation heights of the lines, and the horizontal velocity vectors.

A frame of the synthesised images, heights and speeds are shown in fig. 5.3. In this case, the 3mm surface is so highly corrugated that the horizontal divergence and vorticity fields are somewhat meaningless as they do not correspond to any well defined surface except in the few instances surrounding vortices. For this reason, the speed is depicted instead of the gradients that were typically shown in chapter 2. The figure highlights that the typical nature of the chromosphere is apparently random with little structure. Where swirling structures are present, the heights of features becomes notably coherent and plasma motions are consistently faster in these locations.

The first version of DV used for recovering heights, which for this piece is dubbed DeepVel-heights (DV_h) used only single frames of the sub-image pairs, with the 3mm intensities as inputs and the heights as the labelled outputs. The second network, dubbed DeepVel-chromosphere (DV_c), was trained using the two time-consecutive images and heights from simulation as the input, and the horizontal velocities as the output variables. Both networks converged within 50 epochs. No architectural changes were performed in order to test how well DeepVel could adapt to new scenarios with its out-of-the-box structure. As such, the choice of hyperparameters remains the same as in chapter 2.

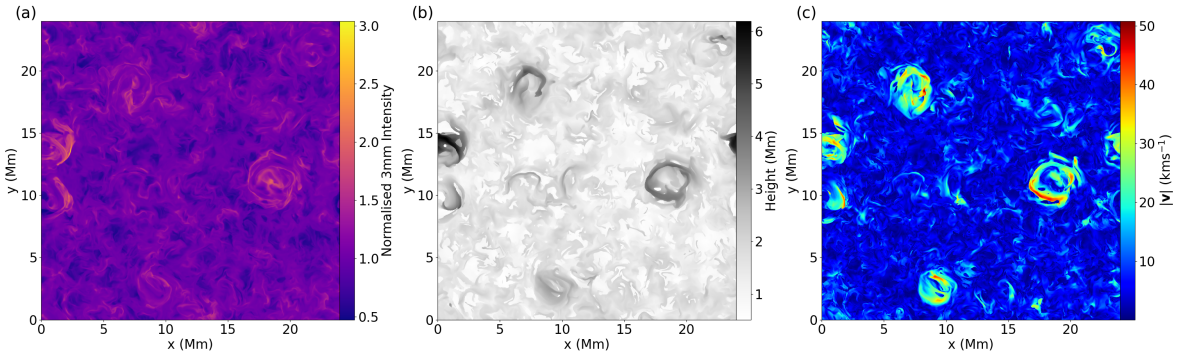

Figure 5.3: *Example of (a) an image synthesised in 3mm band of the ALMA radio telescope, from the Bifrost simulation (ch024031), (b) the radiation formation heights corresponding to the image and (c) the magnitude of the horizontal velocity at the formation heights.*

Whilst these networks were used as a cascade network, feeding into each other, for simplicity their training was performed independently in order to benchmark the respective performance of each network and set out what may be achievable. That is, $\mathrm{DV_c}$ was tested in the ideal case using the calculated heights from the simulation. Its performance was then also tested using the recovered heights and the performance of these two outputs were compared in order to fully understand the current state of its applicability to the real Sun. The results of these tests are shown in the following subsection.

5.2.3 Vortex Detection Methods

The λ_2 criterion is an Eulerian method for detecting swirling structures in flows. It identifies the vortex cores by placing a restriction on the eigenvalues of a tensor which determines the shear and rotation of a velocity field. It works by defining the gradient tensor of the velocity field $\nabla\mathbf{v}$ and decomposing it into a symmetric rate-of-strain tensor $\mathbf{S} = \nabla\mathbf{v} + (\nabla\mathbf{v})^{\mathrm{T}}$ and an antisymmetric vorticity tensor $\mathbf{\Omega} = \nabla\mathbf{v} - (\nabla\mathbf{v})^{\mathrm{T}}$, from which the symmetric tensor $\mathbf{M} = \mathbf{S}^2 + \mathbf{\Omega}^2$ is defined. A condition on the second eigenvalue of $\mathbf{M}$ being negative ($\lambda_2 < 0$) defines a vortex core. This is because 3-dimensional flows have 3 associated eigenvalues with the symmetric tensor, which are put into descending order, if the second of these values is negative, then there is a locally decreasing pressure in at least two of the principal direc-

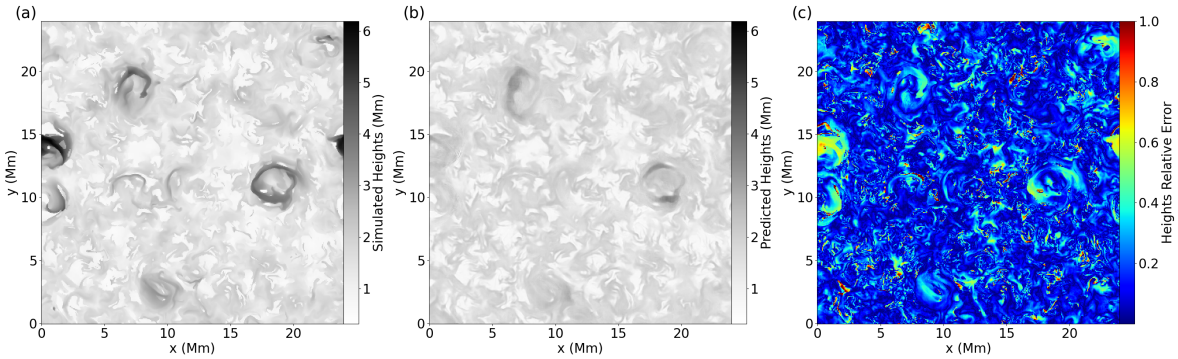

Figure 5.4: *The (a) radiation formation heights of a frame of the Bifrost simulation highlighted in fig. 5.3, (b) heights recovered by the version of DV trained to recover heights from 3mm band images of the chromosphere and (c) the relative error of the heights produced by DV.*

tions indicating the presence of a vortex. In two-demensional flows, this simplifies to the requirement that one of the eigenvalues be less than zero. The flow fields were rebinned so that every pixel represents 1Mm, thus revealing only the largest and most coherent swirls, reducing the possibility of turbulent flow regions registering as coherent swirling motions, as well as improving the learning performance of DV by reducing the complexity of the flow fields.

5.3 Results

5.3.1 DeepVel-heights

The network $\mathrm{DV_h}$ was trained on synthesised images of the 3mm band from ALMA, at half of the simulation resolution, to recover the formation heights of the observed radiation. An example frame from a time in the simulation where swirling features are present is shown in fig. 5.4. A number of metrics are presented for the success of the two networks, these metrics are highlighted in the methods of chapter 2.

As seen in fig. 5.4 the measured radiation is generally captured in a range of heights which are typically associated with the chromosphere (1–2Mm) above the surface). In a handful of locations, where there is the presence of coherent vortical motions, the radiation forms at

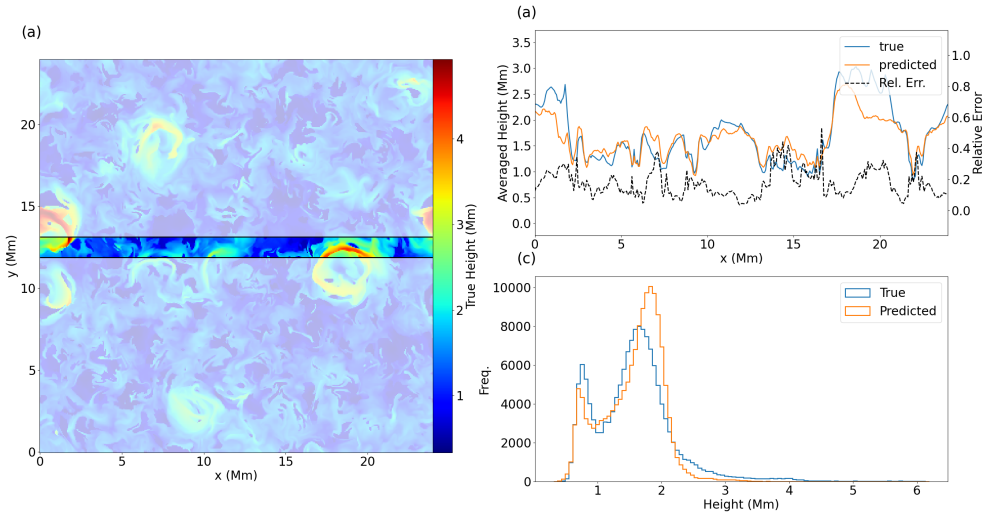

Figure 5.5: *The (a) heights over the field of view of a frame from the Bifrost simulation. The highlighted regions depicts a slice containing swirling features, where the averaged true heights (blue curve) from simulation and the predicted heights by* $\mathrm{DV_h}$ *(orange curve) are compared on the left-hand axis of panel (b), along with the relative error of the predicted heights (dashed line) on the right-hand axis. Panel (c) depicts a histogram of the height distribution from simulation (blue) and predicted by* $\mathrm{DV_h}$ *(orange) across the entire field of view.*

a much greater height (4–6Mm). This is due to the greater temperature and density of the material trapped within the vortex, increasing the opacity of these features. Panel (b) of the figure shows the recovered heights from $\mathrm{DV_h}$.

Whilst the predicted heights of features are not identical, the shapes of vortical structures are apparently similar and more importantly the vortical structures are still identifiable from the background. In fact, the performance for the height recovery is strong. The relative error for the heights is shown in fig. 5.4(c), in which it is clear that the majority of the heights present in the FOV are recreated within $\lesssim 20\%$ error. To be precise, table 5.1 shows that $\mathrm{DV_h}$ is able to identify heights with an MRE of 13%, the KS test score highlights that the predicted height distribution only deviates by a maximum of 8% from the actual. Additionally, the recovered heights correlate strongly to the synthesised heights with a Pearson correlation of 0.74.

As has already been highlighted in fig. 5.3(c), fig. 5.5(a) shows the radiation formation

heights over a frame from the Bifrost simulation. Figure 5.5(b) highlights how well the heights, averaged over the 1.25Mm (20 pixels) of the slice in the y-direction, are recovered spatially, and fig. 5.5(c) shows how well the distribution of heights is recovered over the entire field of view. Between $x = 0$Mm and $x = 2.5$Mm, and $x = 15$Mm and $x = 20$Mm are swirling structures, highlighted by the significant peaks in height in the blue curve of panel (b). $\mathrm{DV_h}$ recovers the height of features reasonably well with a mean relative error of 17.5%, with relative errors generally increasing inconsistently, over background features that appear at a height of $\approx$ 1Mm. The relative error of height also increases in the regions containing swirls, where there are steep gradients in height and $\mathrm{DV_h}$ struggles to correctly reproduce these sharp changes in the height in and surrounding the swirling structures. The height is overestimated by $\approx 20\%$ over the background features and underestimated by $\approx 12\%$ over the vortical structures. The distribution of recovered heights is shown in fig. 5.5(c). This figure reveals that the recovered heights follow a very similar distribution to the synthesised height, with peaks in approximately the same locations. The only major difference is the second peak is shifted to the right and has a much steeper drop-off and shorter tail as opposed to the true distribution which has a much longer tail for features that exist in the upper regions of the atmosphere. This means that $\mathrm{DV_h}$ does not identify features in the extremes as well as the features that sit within the lower chromosphere. However, this is expected as such high-reaching features correspond to vortical structures, which span a wide range of heights and are only present in roughly half of the available time points. Thus it is anticipated that the network will not see so many of these features and be biased to producing more flows typical of the chromospheric background.

5.3.2 DeepVel-chromosphere

The network for recovering chromospheric velocity fields $\mathrm{DV_c}$ was tested in two scenarios. The first using the synthesised heights and the second using the heights recovered by $\mathrm{DV_h}$. In both cases, the synthesised image were used alongside each set of heights as input.

Figure 5.6 shows the results for both of the cases mentioned above. Panel (e) shows

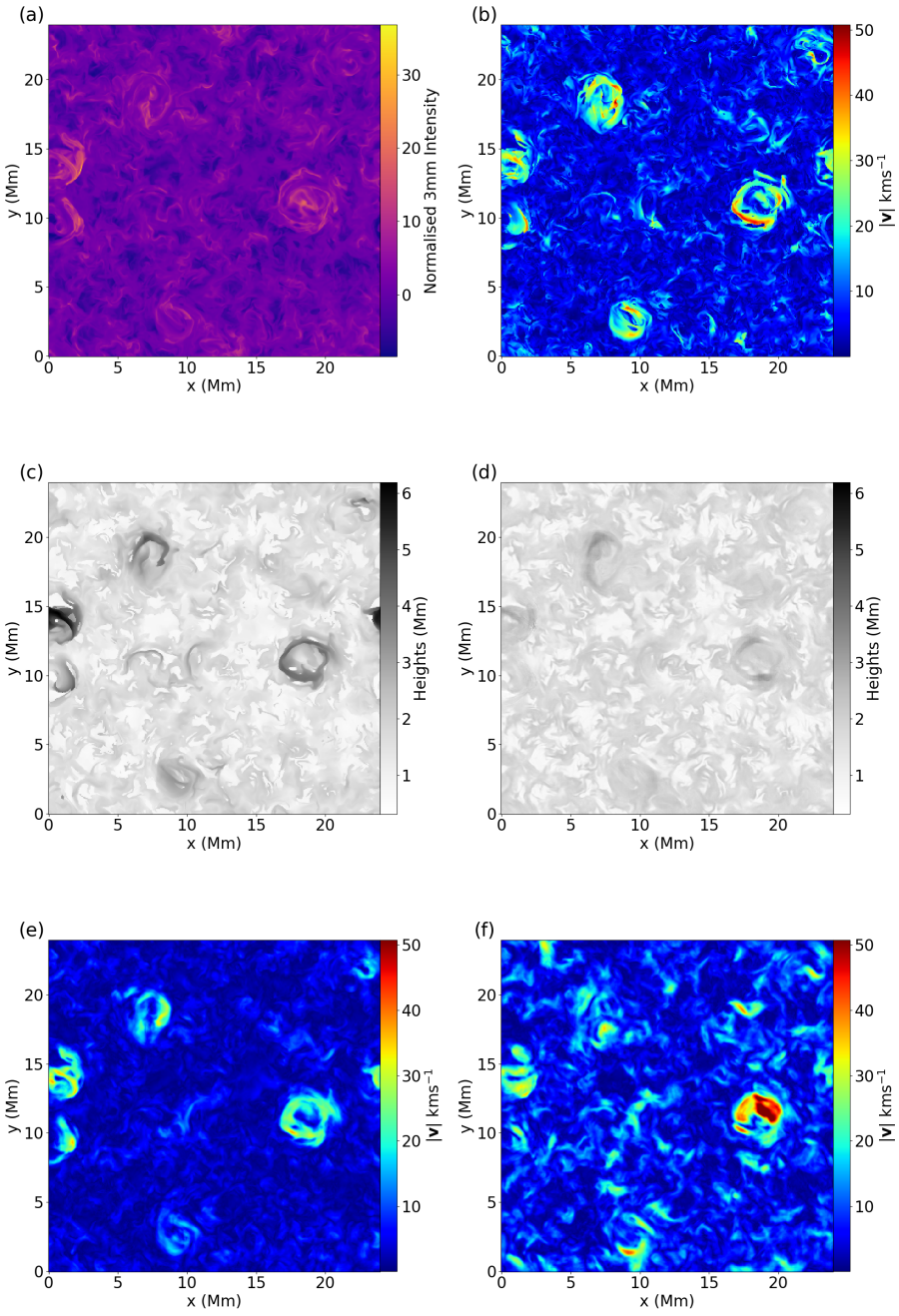

Figure 5.6: *The (a) 3mm intensity of simulated chromosphere from Bifrost, normalised by the mean intensity, (b) the true values of the horizontal speed at the heights of the 3mm band line formation, (c) the heights of the 3mm band line formation, (d) the heights predicted by DV of the 3mm band line formation (see fig. 5.4), (e) the velocities predicted by DV using both the true heights and (f) the predicted heights*

the output by combining the inputs (a) and (c) (the true heights) and panel (f) shows the output produced by combining (a) and (d) (the predicted heights). Panel (b) shows the target velocity field.

Addressing the first case, in which the true heights of features are used to predict the velocity field, we see that the predicted flows are typically slower than the actual flows. In the background flows, flows are underestimated to have around half of the actual speed and a only some of the fine scale structures are identified by $\mathrm{DV_c}$, i.e. the flows from simulation have lots of small-scale flows which move with 2–3 the speed of the background. Larger scale structures such as the vortices are apparent in the flows recovered using the true heights. Their strength is not exact but has the correct order of magnitude. Furthermore, fig. 5.7 shows an example of the a recovered vortex's direction of rotation. Not only are vortical flows detected by $\mathrm{DV_c}$ but they are also oriented with the correct direction of rotation.

In the case where the heights predicted by $\mathrm{DV_h}$ were used as inputs, the predicted velocity field contains a good deal of noise, as anticipated due inputs containing error produced by first network. In particular, the background flows contain small-scale features, but these do not correspond with any of the flows seen in the simulated velocity fields. Most of the large scale swirling structures do appear present in the recovered flows, but these are harder to distinguish against the background.

The results presented in table 5.1 highlight the performance of DV_c in both cases. Namely, it is shown that the correlations of the velocities, when using the true values for the heights of radiation, are strong (> 0.5), the correlation between the true and predicted flows becomes weaker when using predicted heights. The derivatives of velocities over the FOV, which all share a weak correlation with those from simulation. The KS test statistic shows that distributions of the horizontal velocity components differ by a maximum of $\approx 12\%$ when using the true heights and by $\approx 15\%$ when using the predicted heights. Thus, this suggests that $\mathrm{DV_c}$ has learnt to replicated much of the velocity field distribution. The MRE is high, however this has two major contributing factors: the orientation and the magnitude of velocities. From fig. 5.6, the speeds of flows are generally captured well over the turbulent background, but

		PC	KS	MRE	NDP
Height		0.74	0.08	0.13	-
$\mathbf{v}$	TH	-	-	-	0.48
	PH	-	-	-	0.31
v_x	TH	0.66	0.13	0.83	-
	PH	0.36	0.18	1.14	-
v_y	TH	0.64	0.12	0.81	-
	PH	0.36	0.12	1.21	-
$\nabla \cdot \mathbf{v}$	TH	0.19	0.17	1.01	-
	PH	0.11	0.09	1.13	-
$\nabla \times \mathbf{v}$	TH	0.19	0.17	1.01	-
	PH	0.11	0.09	1.13	-

Table 5.1: *Metrics showing the performance of the* $\mathrm{DV_h}$ *and* $\mathrm{DV_c}$ *networks.* $\mathrm{DV_c}$ *was tested in two cases, one where the true radiation formation heights (TH) were used for inputs and one where the predicted heights (PH) from* DV_h *were used.*

the orientation less so; the converse is true over regions containing coherent swirls where the orientation is captured well (see fig. 5.7) but the speeds are under/overestimated depending on the heights used for the input. The velocities are not as well aligned as those recovered in the photosphere (see chapter 2), they are misaligned by at least an average angle of $\approx 60°$.

The less-than-ideal performance of $\mathrm{DV_c}$ is mostly due to the dominance of the background flows, where motions are mostly incoherent. By restricting attention to only higher speed flows ($> 20\mathrm{kms}^{-1}$), the performance notably increases. Flow metrics were compared in regions where high-speed flows were present in the simulation as these correspond strongly with coherent structures such as vortices. An example of such a region is shown in fig. 5.7. Visually, this figure highlights that the flows, at least in the case of using the radiation formation heights from simulation, shown in panel (b), retain much of the shape of the original vortex, and miss only fine-scale structure which is present on the sub-Mm scale. The direction and magnitude of the flows within the vortex are aligned well, only the much greater speeds $\gtrsim 40\mathrm{kms}^{-1}$ are not identified so well. In addition to this, the PC over these structures increases to 0.83 and the NDP improves to 0.82 (corresponding to an average difference of $\approx 35°$). These values improve slightly for the flows recovered using the predicted heights to

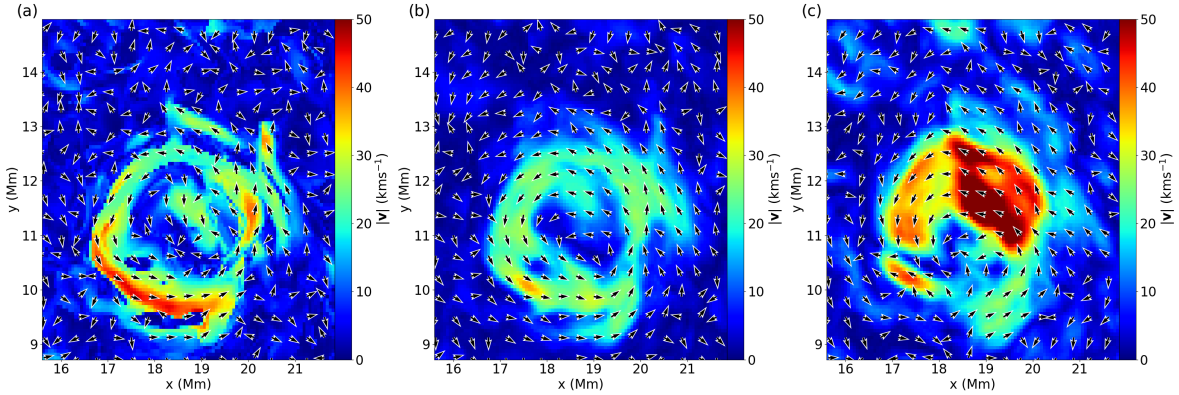

Figure 5.7: *The speed (shown by the colouration) and orientation of the same swirling structure from (a) the Bifrost simulation, (b) recovered by* $\mathrm{DV_h}$ *using a synthesised 3mm band image and the radiation formation heights from simulation, and (c) the same as (b) but using heights predicted by* DV_h.

a PC of 0.39 and NDP of 0.462 (difference of 62.5°).

5.3.3 Validation with the λ_2-Criterion

The λ_2 method for vortex detection was deployed on velocity fields recovered by DV. The results of this are shown in fig. 5.8. The vortices in the frame shown persist on the order of 10min and in this particular case 11 vortices in the simulation were detected by the method. The velocity field produced by $\mathrm{DV_c}$ using the heights from the simulation only presents 6 vortex cores, whereas when the predicted heights were used the method is able to identify 11 also. Of the 11 detected vortex cores found using the true heights, $\mathrm{DV_c}$ was able to identify the location for 3, exactly. When using the predicted heights, only 2 were identified in the correct location. However, the location of the detected vortex cores are all close (within 2Mm) to the locations of the strongest swirls present in the simulation, aside from one instance where it is likely flows in the noisy background falsely generated a rotational motion. Additionally, all of the vortex cores reported from the recovered flow fields sit within the boundary of a true vortex with the singular exception.

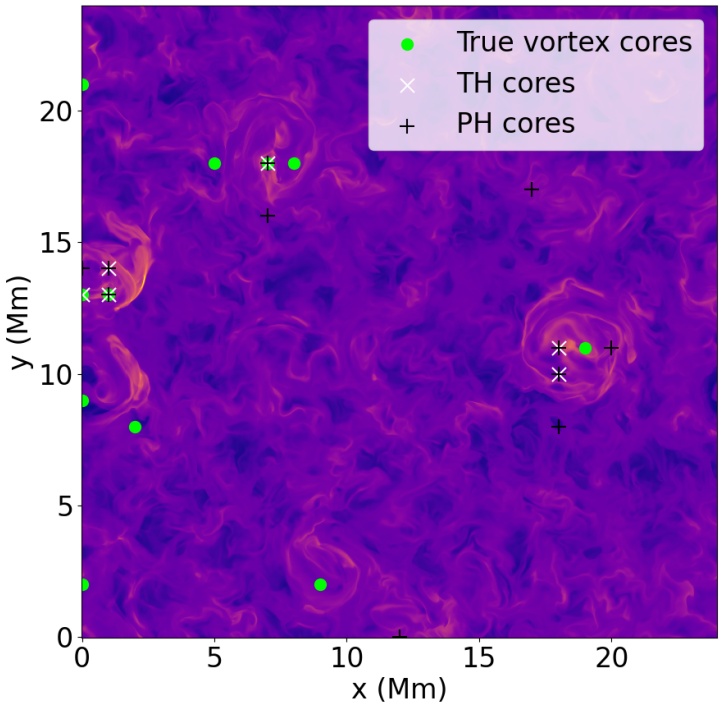

Figure 5.8: *Vortex cores detected in the simulation flows as well as in flows recovered by DV using both the true formation heights and predicted formation heights superposed on the 3mm intensities.*

5.4 Summary

In this chapter, the first step toward being able to recover flows by means of using simulated data over a coronal hole in LTE conditions has been evaluated, by training the DV network on synthetic images in the 3mm band observed by ALMA. The synthetic 3mm images cover a wide range of heights and share resolutions with modern telescopes such as SST CRISP and CHROMIS (which view the chromosphere in Hα and Ca II k lines, respectively; see Scharmer et al. 2002), GREGOR (see Schmidt et al. 2012) and IRIS (see De Pontieu et al. 2014). Whilst this is not a realistic setting for validating DV, and the training data does not contain all of the qualities of observational images such as atmospheric seeing, point spread functions (PSFs) and noise, these first results do highlight a promising avenue for recovering

chromospheric velocity fields.

In this chapter, two versions of DV were trained: DV_c, for recovering chromospheric velocity fields from 3mm band images and the radiation formation heights, and DV_h for recovering the radiation formation heights from the 3mm band images. Both networks showed success at determining features that correlate well with the true fields produce by the Bifrost simulation.

DV_h was able to reproduce the heights of features well with a strong PC and a difference in distribution of 12% and a mean relative error of 17% over the entire field of view, showing that the NN is able to recover heights reasonably, which come from a closely matched distribution. The only significant errors in height were identified where the heights reach the extremes of the true distribution's tails. This is somewhat expected as the extreme heights are present almost exclusively within swirling structures and these take up only a small portion of the data available for training. Thus the NN will be biased to produce values that fit within typical parameters. On the other hand the DV_c shows promising performance in identifying velocities in regions containing more distinct structures, however, does not perform as strongly at identifying the velocities of background flows. In this instance, the NN is able to determine the difference between the background and the vortical structures of interest, identifying most of the large-scale vortices present in the simulation as well as correctly identifying their direction or rotation. This work is a key step forward in flow recovery in the chromosphere as the coherent structures of interest have significant implications for energy transport in the chromosphere.

Despite the novelty of the methods presented in this chapter, the results indicate two major areas for improvement. Firstly, if this method is to be deployed in the future, the performance of DV_h for predicting heights needs to be improved to optimise flow recovery. Secondly, the recovery of flows within swirling structures must be enhanced to more precisely identify swirl locations and characterise their vorticity and topology.

The synthetic chromospheric images used for training contain many incoherent, transient features at sub-Mm scales alongside coherent, persistent vortices at Mm scales and up. This

multiscale mixture produces noisy signals that appear across a range of heights and complicates learning when the network should focus on the coherent swirls of interest. One route to address both height and vortex-recovery issues is to better separate the features of interest from background noise, either by preprocessing or by changing the architecture. For preprocessing, vortices can be located in the simulations using a vortex-detection algorithm; the resulting vortex masks could be provided as an additional output target or used to weight the loss function so the network emphasises vortex regions. Alternatively, modal filtering, e.g. proper orthogonal decomposition, can suppress incoherent background features prior to training and this same step can be used on new inputs. However, these both require computation time, slowing the overall pipeline.

For architectural changes, the current DeepVel variant was optimised for dense plasma regions with coherent motion; a more suitable design could be a multi-scale U-Net for segmeting vortices from the flow, improving feature extraction similar to the U-Net designed by Tremblay & Attie (2020) for ARs. Another approach to better capturing persisting structures is to exploit the temporal coherence of vortices by extending the input to a longer time series and employing a convolutional long short-term memory (ConvLSTM; see Shi et al. 2015) network. ConvLSTMs use recurrent connections to construct temporal relationships between features as well as convolutions for identifying spatial patterns, which should support the network in distinguishing persistent vortices from short-lived noise. Finally, errors in regions with sharp gradients can be mitigated by activation functions that better capture high spatial frequencies, e.g. sinusoidal representation networks (SIRENs; see Sitzmann et al. 2020), and by modifying the loss function to penalise gradient errors and incorporate physics-informed terms such as vorticity. Furthermore, the general performance of the cascade could be improved somewhat by combining their training process. By using a top down approach where $\mathrm{DV_c}$ is trained to recover the velocities from simulated heights first, and then training DV_h to recover the heights based on the quality of the predicted velocity fields, the performance could improve significantly (Zhang et al. 2021). However, this may have unexpected effects for $\mathrm{DV_h}$ as it would not be learning to minimise the loss for recovering heights. The network

for recovering heights is useful on its own for recovering additional in formation from solar imagery.

Chapter 6

Discussion and Conclusions

The matter of this thesis has address the applicability of a modern machine learning approach for the recovery of plasma flows in the solar atmosphere by validating its success and studying new possible outlooks with cutting edge simulations. As computational modelling and observational instruments improve, there is a vast wealth of new data, which needs to be analysed and utilised effectively. Neural networks provide a fast computational approach to dealing with these data as they not only provide access to new data, which is otherwise irretrievable at the scales and speeds shown in this report, but also give rise to new insights due to the quality of the data, as has been presented in this work. This thesis reports the practicality of using DV in some of the most complex solar regions and deals with understanding how well it can generalise to new data. In addition, new methods are presented for the validation of such flow recovery and how CNNs can be applied in observational contexts whilst adapting it to entirely new scenarios representative of non-LTE dynamics, for which it was not intentionally designed to deal with.

Chapter 1 highlights that the Sun is a complex and highly stratified object, in which there are many concerning problems relating to the motions of plasma, namely the dynamics leading to the onset of solar flares and energy transport through magnetised plasmas. The problem of measuring and studying plasma flows in magnetically dominated regions is addressed in chapter 2, and comparisons were made against the current widely used method

FLCT. Once establishing that DV is successful in recovering flows, moreso than an existing method, the topology of recovered flows were derived by means of identifying Lagrangian coherent structures in chapter 3. The LCSs reveal insight to skeleton for the flow dynamics in the photosphere and therefore evidence DVs ability to generalise well to data from new simulations and produce high quality flow fields across multiple scales. With the photospheric capabilities of DV established, chapter 4 shifts the focus to a use case of recovered flow maps, using changes coherent structures as a detection method for the presence of ARs with notable success, consistently identifying regions of $>$ kG levels of magnetic flux density at the (meso)granular scale. Chapter 5 extends the use cases of DV into the chromosphere by training on synthetic images from the Bifrost code and shows success in the way of identifying coherent vortices.

It was found that DVs performance is unrivalled by the popular FLCT method at short time scales, both in efficiency and accuracy supporting a strong correlation and alignment with the true flows even when faced with unseen data from a second simulation. Additionally, DV flows are accurate enough to reproduce key repelling and attracting coherent structures that govern the flow throughout the evolution of a pore. This has two major implications for the use of NNs for flow recovery in solar physics. First, in the absence of any ground truth data for solar flows, cross-testing the network on other simulations establishes how well the network is expected to perform on real-world observations. Thus predicting flows with the levels of accuracy seen for the first time in chapter 2 promotes a promising potential for the application of NNs for predicting flows from real data containing ARs and pores. Secondly, velocities are predicted with enough confidence to reproduce LCSs, seen in chapter 3. Thus, the flows that are identified by DV are physically consistent with the simulation, exhibiting the same flow dynamics that are present, highlighting that unlike other methods like FLCT that are consistent with only an advection equation (Schuck 2005), DV is able to emulate flow physics correctly. Where previous studies highlighted DV performance by testing only on the training simulations and with image properties that both mimic the SDO/HMI instrument, supporting the preprocessing steps for FLCT (e.g., see Tremblay & Attie 2020, Tremblay, Benoit et al.

2021), the work presented in this thesis indicate more strongly how error that is generated by the network propagates through a time series of velocity fields, without additional spatio-temporal smoothing which thereby removes them in the first place. However, the results of the analysis performed in this report do not address observations directly and only aim to understand how the network adapts to new data. The results reported by Tremblay et al. (2018) are similar enough to indicate that performance would be generally unaffected by telescope optics or the presence of 5-minute oscillations. However, effects such as the curvature of Sun or the drifting of features and differential rotation remain yet to be studied in the context of using DV. In addition, atmospheric seeing poses a further challenge for ground-based telescopes with varying degrees to which it affects imagery (Padinhatteeri et al. 2010). In order to mitigate this a number of blind deconvolution methods have been proposed for removing the effects of turbulence in the atmosphere (e.g. Liu et al. 2022). An advantage of using DV in observational contexts is that provided a high enough resolution and cadence data for in the training set, the data can be manipulated and convolved with effects to suit any desired instrument so long as the optical properties of the telescope are known and hence removing some of the difficulty in transitioning the network to work with real-world examples.

The ability for DV to accurately determine LCSs in flows was shown to be useful in chapter 4, where evidence was presented that changes in granular-sized repelling structures can be used as an indicator for the presence of high concentrations of magnetic flux at granular to mesogranular scales. From this, results reported by the FTLE fields derived from recovered flows by DV showed almost the same level of success as a classifier for the presence of ARs as those derived from simulated velocities. Combined with the fact that simulation flows have been shown to be able to forecast an emerging AR up to 10 hours prior to its appearance in the photosphere, there is now significant evidence that DV could be deployed for applying such forecasting methods for the analysis of ARs from observational data.

In addition, the robustness of NN flow recovery was established further by training DV on images of radiation synthesised by a non-LTE code, that was representative of the chromosphere. Whilst the method requires more refining in its current state before it can be

determined to be useful in a true solar context, the network was able to identify coherent vortical motions, detected by the λ_2 criterion, against a turbulent background with promising success in the ideal case. This was in tandem to training a new version of DV which was able to retrieve the formation heights of features with high levels of accuracy enabling the study of the 3D structure of chromospheric environments.

In the theme of this thesis, the ability for CNNs such as DV to recover velocities and identify flow features present in synthesised images was identified to be strong. DV was able to identify flows at the subgranular scales with a great success in photospheric conditions. The highly accurate flows could be used to determine precise LCSs present across simulations, indicating that DV is a robust method for predicting horizontal flows from new data which can be used for forecasting and identifying ARs. Furthermore, the DV network was shown to be suitable for predicting flows outside of its intended use, being able to identify coherent vortices in flows close to the transition region. These results highlight that even though the application of NNs is still relatively new in the fields of solar physics and space weather, it already opens up a realm of possibilities for studying the Sun in never-before-seen detail, shortening the gap between numerical modelling and observational studies.

Chapter 7

Future Works

1. **Observational Testing**

   The application of the version of DV trained in this thesis is yet to be seen in the context of real solar imagery. Convolving training images with the PSF of a number of satellites and comparing results with those seen through inversion techniques from helioseismology is a crucial final step in ensuring that the DV network is reliable when transitioning to working with observational data since there is no ground truth for measuring horizontal velocity fields. Another approach to validating the network capabilities is to apply the methods shown in chapter 4 to determine if the forecasting and tracking capabilities of DV extend to observations. A number of steps can be made in understanding the influence of observation effects on results such as training with simulations 5-minute oscillations (p-modes), convolving simulation data with telescope optics and introducing noise into the testing set to mimic the effects of turbulence in the Earth's atmosphere. The initial stages of this testing is highlighted in appendices A and B.

   An additional point linked to this is for the versions of DV presented in chapter 5. These are currently trained on data which does not inherently match the seeing properties of any existing telescope. Whilst the 3mm band is observed by ALMA, the resolution of this radio telescope is significantly lower than the resolutions of the data that DV was

trained on. Despite the low resolution of physical observations from ALMA, there is an advantage to studying the Sun in the 3mm band: longer wavelengths are less susceptible to atmospheric seeing. For example the Fried parameter, which increases with the wavelength, is defined such that for a telescope with an aperture diameter smaller than the Fried parameter, is then diffraction limited and observations are less influenced by seeing (e.g. see Irbah, A. et al. 2016). The parameter describes the influence of atmospheric seeing for a given wavelength, such that if the diameter of the aperture is less than the Fried parameter. Thus ALMA is less affected by atmospheric turbulence than other similar instruments that see in shorter wavelengths. By synthesising other spectral lines such as Hα or the Ca-II, seen by SST for example, will be much more indicative of the NNs performance on the true Sun where plenty of data exists in these spectral lines.

2. **Training and Architecture Changes**

   An almost endless number of changes could be made to the DeepVel network. However, there are some notable obvious steps that should be taken for increasing the performance of the network in the contexts provided in this thesis. First and foremost would be to modify the network so that it becomes physics informed. That is alter the loss function so that it at least includes the induction equation, thus the network will learn to produce velocities that minimise the residuals of the governing equations and hence produce flows that are consistent with the physical system and not just identify patterns from simulation flows. An additional change to the architecture would be to change the activation function so that it performs better at producing gradients that are more consistent with flows observed in simulation, one approach to this is the SIREN activation function (Sitzmann et al. 2020), which were designed to deal with networks that produce signals as the solution to differential equations.

   For the version of DV presented in chapter 5, a full cascade network should be trained and tested to determine whether the output velocities from the top network can be optimised further by combining the training process. That is, the training for the $\mathrm{DV_h}$

should be dependent on the losses calculated in DV_c.

In addition to changing the network architecture, a more expensive yet simpler solution is to make changes to the training set. Of which there are two straight forward approaches: the first being to train DV on a range of simulations, so that a full range of physics and effects can be seen by DV; the second approach is to train a series of networks in an ensemble approach (e.g., see Ganaie et al. 2021) where a number of outputs from networks are combined to produce a coherent output.

3. **Lagrangian Vortex Detection**

   The current results of chapter 5 are not so robust and currently provide a proof of concept for applying DV to further regions of the solar atmosphere, outside of the photosphere. In the interest of identifying coherent flow patterns from otherwise noisy and turbulent images would be to apply Lagrangian techniques such as LAVD (Haller et al. 2016) in order to further understand and improve the chromospheric version of DV. In the interest of detecting locations of coherent structures and specific parameters, the network could also be trained to minimise residuals of parameters such as swirling strength and vorticity.

Appendix A

Point Spread Functions

Images of the solar atmosphere are subject to optical aberrations due to the lensing of light by the instrument and obstructions due to atmospheric seeing (see Padinhatteeri et al. 2010). This optical smearing of light can be modelled using a point spread function (PSF). The PSF is described in two parts, a central Gaussian core and the tail of the curve. Mathematically this is given by

$$\mathrm{PSF}(r) = (1-\epsilon)e^{-\left(\frac{r}{w}\right)^2} + \frac{\epsilon}{1+\left(\frac{r}{W}\right)^{\kappa}}, \tag{A.1}$$

where ϵ determines the relative size of the terms, r is the radial distance across, w describes the rate at which the Gaussian drops off by $1/e$, W describes the half-width half-maximum of the tail and κ determines the rate at which the tail drops off (see Wachter et al. 2012). As well as these, instruments will typically pick up noise on their sensors.

Synthetic intensity images from the MURaM simulation were convolved with the PSF of the SUNRISE/IMAX instrument (see Martínez Pillet et al. 2011). This application of the PSF helps to determine the influence of optical effects on $\mathrm{DV}_{\mathrm{R2D2}}$ without specific retraining to deal with them. IMAX presents finer scales (see Asensio Ramos, A. et al. 2017, for details) than what are available in within simulations thus, degrading images to closely match the training data will allow us to determine the success of the network at the some of the finest

available resolutions.

The IMAX instrument observed the photosphere at 525.02nm (the Fe I line). The Fe I line corresponds to temperatures of around 4000-6000K, which relates to heights of the photosphere to the lower chromosphere. IMAX had an effective resolution of ≈ 0.11" (equating to ≈ 0.055"pix^{-1}) and has a cadence of $\approx$ 33s. The telescope has an aperture of 1m. The instrument took high resolution images of the photosphere from high in the atmosphere and thus circumvents most atmospheric effects. Additionally, the telescope has a wide aperture and relatively small FOV, meaning that the light entering the instrument will be weakly distorted by the lens, as opposed to other instruments which have smaller apertures or much wider FOVs.

During the training process, noise is already added to images, this is done to reduce the potentiality of overfitting the network. However, this will also somewhat simulate the effects of noise in observations. With the large number of high-resolution space-based telescopes, we argue that DV may be readily deployed for real-world observations by deconvolving the PSF from observations, or by retraining with the PSF of instruments for uses specific to an instrument.

In order to simulate the instrument the PSF was calculated and convolved with an image from the MURaM simulation. This was achieved by selecting a region which covers the same angular distance as those taken by IMAX, upscaling the image using linear interpolation to match the resolution of the instrument (MURaM has a pixel resolution of 96kmpix^{-1}, opposed to the 39.9kmpix^{-1} resolution of IMAX). The PSF was identified to have a FWHM of $\approx$ 80km, since this is smaller than the grid size of the MURaM images it is expected that almost all of the structure in the image at the 96kmpix^{-1} resolution of the simulation will be preserved. However, while this establishes that it is not strictly necessary to apply the PSF, for completeness images were upscaled to the resolution of IMAX, the PSF was then convolved with the upscaled images which was then downscaled back to the original resolution of the MURaM images. Whilst DV is expected to have generalised well enough to recognise features independent of exact resolution, it will perform best on data which most

closely matches the training set, and this process highlights how real-world observations may be transformed for improving performance with the application of DV, for example through up/downscaling with interpolation algorithms. In addition to this, it is expected that some non-physical features will have been introduced to the image during the upscaling process, in effect simulating some additional level of noise.

An example region, PSF and resulting image is presented in Figure A.1. The result of the PSF on the image is a softening of edges and reduction in contrast, especially noticible on small brightening in the image. In other words, the PSF causes light to be scattered across the FOV.

The velocities produced by $\mathrm{DV}_{\mathrm{R2D2}}$, when using the PSF-convolved images, is shown in Figure A.2 in comparison with the true velocities. Whilst the figure shows that $\mathrm{DV}_{\mathrm{R2D2}}$ performs worse at the smaller scales in this case, where some of the fine scale structure within granules has been lost and the transition to the AR is not as clearly defined, the features reproduced by $\mathrm{DV}_{\mathrm{R2D2}}$ are more consistent at the $\geq$ Mm scale. In contrast with testing showing that when the network is applied to raw images from the simulation it produces velocities consistently at the subgranular scale. Despite these differences, the velocity components still support a strong Pearson correlation coefficient of 0.6, highlighting a statistically strong relationship between the recovered flows and the true flows. Similarly, the divergence also supports a strong Pearson correlation of 0.57. The PSF also results in a normalised dot product with mean 0.46 (corresponding to an angle of $\approx 60°$) and median 0.76 ($\approx 40°$), that is half of the flows recovered are generally oriented in the same direction as the simulated flows.

Whilst the performance of $\mathrm{DV}_{\mathrm{R2D2}}$ is affected by the inclusion of the PSF in images (as well as potentially from the introduction of noise through upscaling and downscaling images), the performance is still strong and presents the potentiality of being used with observations for identifying velocities the will provide some agreement with the true flow field. It is worth noting that this version of DV was not trained to deal with telescope optics or any significant noise within the data. It is therefore reasonable to suggest that the performance of DV would

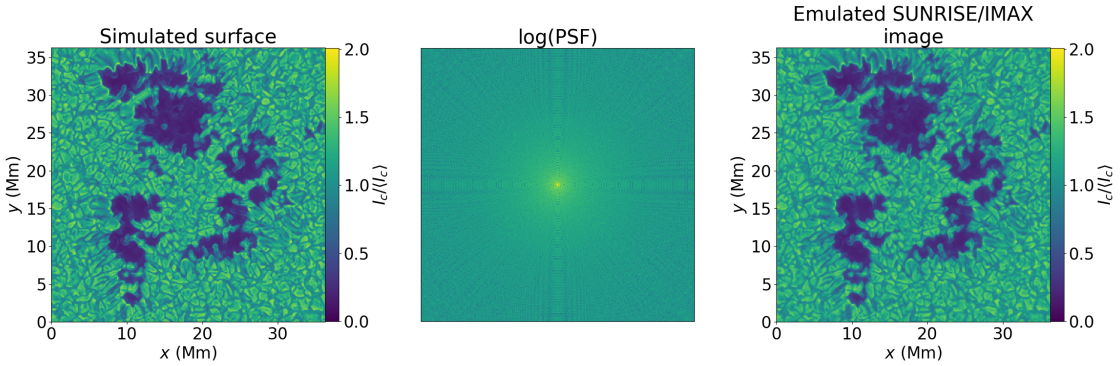

Figure A.1: *This figure depicts (a) Original synthesised image from simulation, (b) log plot of PSF to show the effect of the lens and (c) the resulting image synthesised to look like one from the IMAX telescope after convolving with its PSF.*

be expected to increase if the training set images were to be convolved with telescope optics and the network retrained, or by deconvolving the PSF from real-world observations.

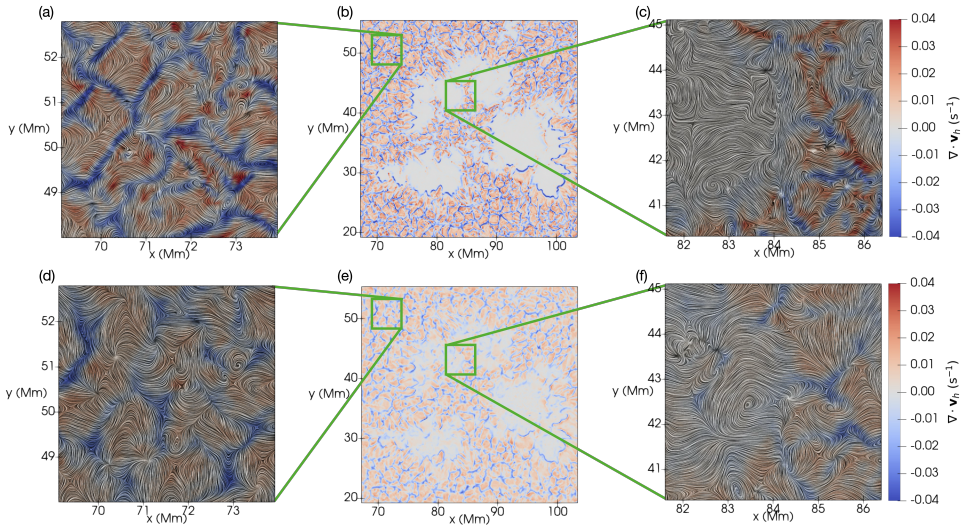

Figure A.2: *This figure depicts the divergence field over a part of the MURaM simulation (b) and the LIC visualisation of the flows, superposed on the divergence in a region with* $<$ kG *magnetic flux (a), and a region with* $>$ kG *magnetic flux (c). The same region as (b) is shown for velocity fields produced by DV when using the images convolved with the IMAX PSF in (e). The same* $<$ kG *and* $>$ kG *regions, described above, are also shown in* (d) *and* (f)*, using the DV-recovered velocities from images convolved with the IMAX PSF.*

Appendix B

P-modes

In order to cross-validate the version of DV presented in chapter 2 in the context of the solar atmosphere, flows from the QS Stagger simulation (see Stein & Nordlund 2012, Stein et al. 2024) were recovered. The Stagger data presents a model of QS magnetoconvection in a shallow box, where a weak background magnetic field is advected by plasma and forms small concentrated regions of magnetic flux. Continuum intensity images, with the same 96kmpix^{-1} spatial resolution and 120s cadence as the R2D2 simulation. The key aspect of this simulation, compared to the R2D2 and MURaM simulations used throughout this thesis, is that it contains the presence of short-lived oscillations that are representative of p-modes. These p-modes are oscillations driven by pressure waves within the Sun (see Ulrich 1970).

The results of this testing show the performance of $\mathrm{DV}_{\mathrm{R2D2}}$ on a QS simulation with a very different initial setup to the simulations used for training and validating previously. Previous works from Tremblay et al. (2018), Tremblay, Benoit et al. (2021) already highlight the ability of DV to recover flows from this simulation when trained on the same data, thus providing a benchmark for the performance of DV. In order to apply DV and produce meaningful velocities, the intensities from the synthetic observations were histogram-matched to the ones used for training, i.e. the cumulative frequency distribution of pixel values was matched to those of the R2D2 continuum intensity images. It was shown previously that high-frequency filtering of images provided no real advantage over the default preprocessing

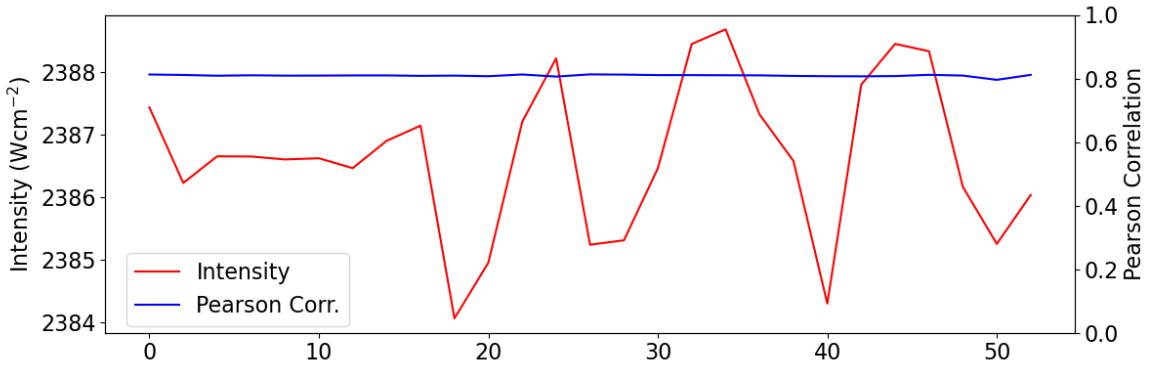

Figure B.1: *Variation of continuum intensity over time against the Pearson correlation of horizontal velocities recovered by DV with the true horizontal velocities from the Stagger simulation. The intensity varies over time due to 5-minute oscillations induced by p-modes. The recovered velocities show negligible variation in their high correlation with the true velocities.*

of images of histogram matching (Tremblay et al. 2018), and therefore similar results were anticipated with the version of DV presented here.

Figure B.1 shows simply how the Pearson correlation coefficient of flows generated by $\mathrm{DV}_{\mathrm{R2D2}}$ varies in time with the 5-minute oscillations. Clearly, the variations in intensity, once filtered by histogram-matching the intensities from Stagger to the QS instensity distribution from R2D2 (as in chapter 2), have no significant effect on the networks ability to identify horizontal flow fields. This also indicates that the p-modes present in the simulation have no discernible effect on the horizontal motions of the flows. Only the vertical/LOS component of the flow would be expected to exhibit any notable difference, however these are not of interest to this study as they are retrievable by means of Doppler measurements.

The velocity fields shown in fig. B.2 show how flows recovered by the NN compare to the true flows from simulation. Almost all of the divergence and converging features were discerned precisely by $\mathrm{DV}_{\mathrm{R2D2}}$ with the magnitudes of both speeds and divergence matching the simulation closely. In addition to this, a number of small scale swirls in the intergranular lanes were identified and resemble those found in the simulation; only the most complex of these flow patterns were not precisely recovered. The majority of the flows present in R2D2 and the Stagger simulation are over regions containing near QS-like granulation, thus the

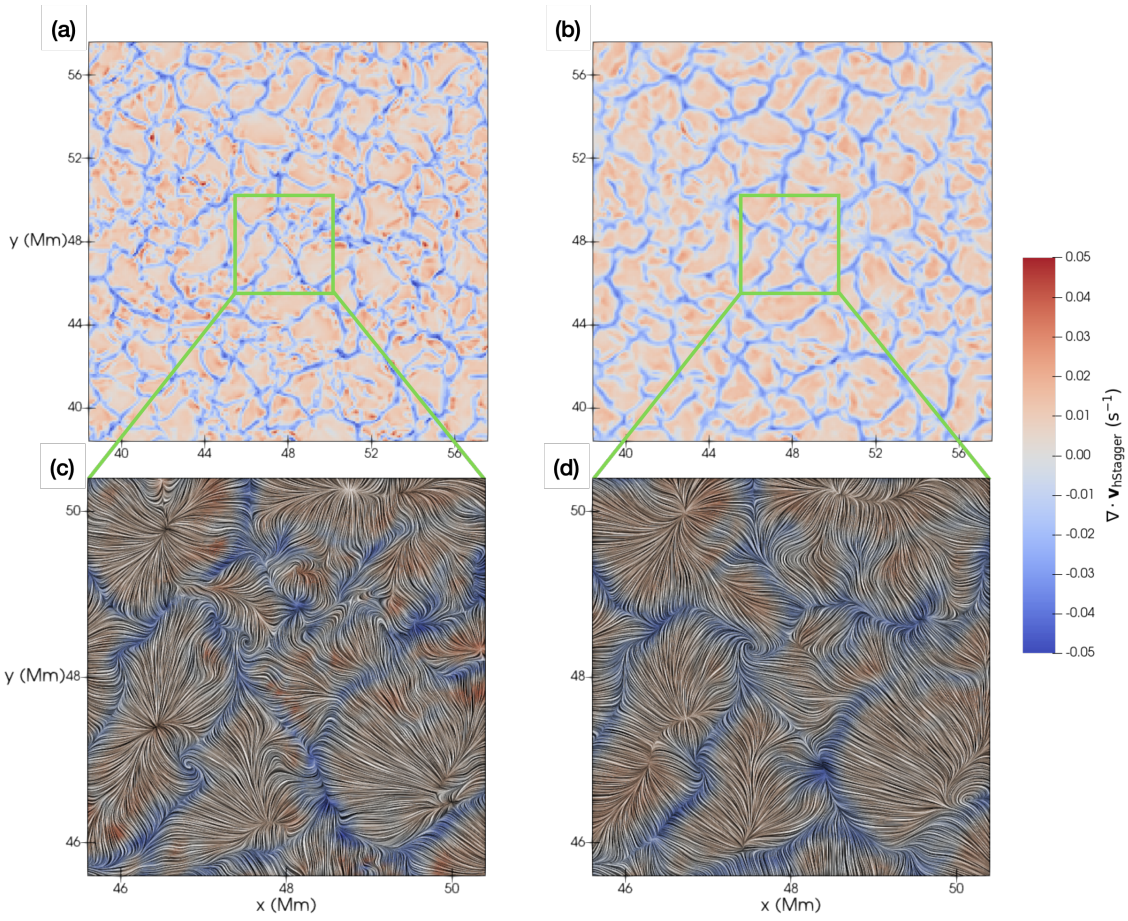

Figure B.2: *Example flows from the Stagger simulation (left column) versus the flows predicted by* $\mathrm{DV}_{\mathrm{R2D2}}$ *(right column). The regions are coloured by the horizontal divergence field and the bottom highlight a region where a LIC field has been superposed on the divergence to highlight the velocity field lines.*

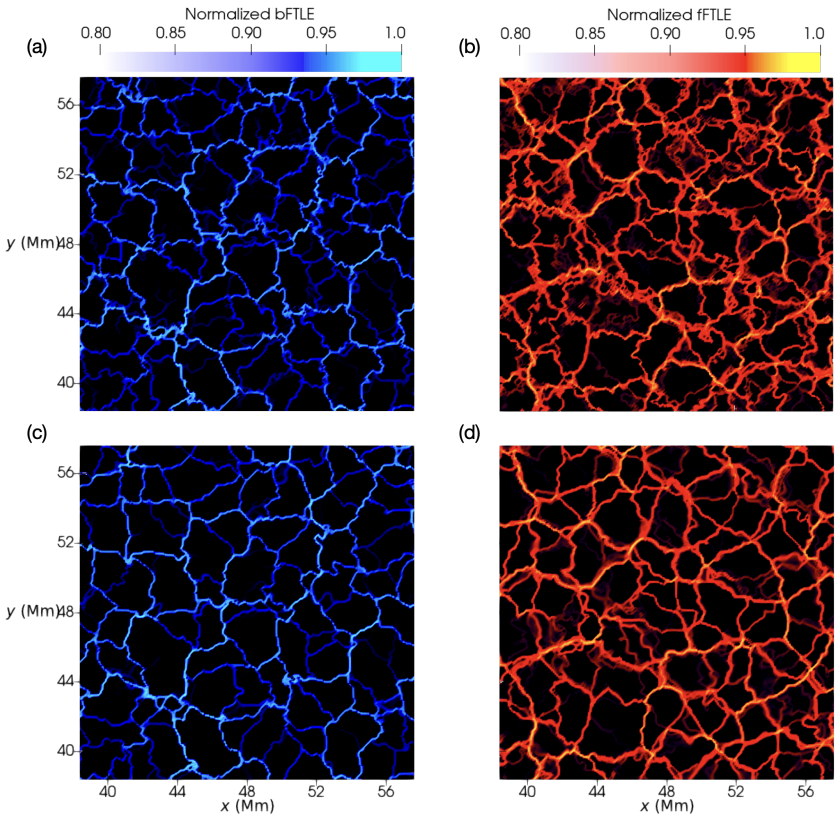

Figure B.3: *The 20-minute backward- (left) and forward- (right) FTLE fields calculated using the flows from the Stagger simulation (top row) and velocities produced by* $\mathrm{DV}_{\mathrm{R2D2}}$ *(bottom row).*

centres of granules where divergence is positive dominate the training data. Therefore the subgranular structures in intergranular lanes that contain complex structures such as saddles and vortices are expected to not be recovered as precisely.

To further highlight the success of $\mathrm{DV}_{\mathrm{R2D2}}$ the 20-minute FTLE fields were calculated over the frame presented in fig. B.2. These FTLE fields are shown in fig. B.3. The backwards-FTLEs shown coalign with integranular lanes and show the attracting downflows that are present surrounding the granules. The forward-FTLE field highlights the repelling structures which cross over the centres of granules. Panels similarities between the fields for both Stagger and NN-recovered flows are highly similar and only significantly differ in regions where weak transport barriers are present. These results strongly align with what was previously

established in chapters 2 and 3, that DV is able to recover LCSs well across simulations and therefore capture the flow dynamics present, even in complex regions.

Bibliography

Agarap, A. F. (2019), 'Deep learning using rectified linear units (relu)'.
URL: *https://arxiv.org/abs/1803.08375*

Alfvén, H. (1942), 'Existence of electromagnetic-hydrodynamic waves', *Nature* **150**(3805), 405–406.
URL: *https://doi.org/10.1038/150405d0*

Asensio Ramos, A., Requerey, I. S. & Vitas, N. (2017), 'Deepvel: Deep learning for the estimation of horizontal velocities at the solar surface', *A&A* **604**, A11.
URL: *https://doi.org/10.1051/0004-6361/201730783*

Babcock, H. W. (1967), 'The zeeman effect in astrophysics', *Physica* **33**(1), 102–121.
URL: *https://www.sciencedirect.com/science/article/pii/0031891467902637*

Bahng, J. & Schwarzschild, M. (1961), 'Lifetime of Solar Granules.', *The Astrophysical Journal* **134**, 312.

Beeck, B., Collet, R., Steffen, M., Asplund, M., Cameron, R. H., Freytag, B., Hayek, W., Ludwig, H.-G. & Schüssler, M. (2012), 'Simulations of the solar near-surface layers with the co5bold, muram, and stagger codes', *A&A* **539**, A121.
URL: *https://doi.org/10.1051/0004-6361/201118252*

Bell, B. & Meltzer, A. (1959), 'The Doppler Widths of Solar Absorption Lines', *Smithsonian Contributions to Astrophysics* **3**, 39.

Birch, A. C., Braun, D. C., Leka, K. D., Barnes, G. & Javornik, B. (2013), 'Helioseismology of Pre-emerging Active Regions. II. Average Emergence Properties', *The Astrophysical Journal* **762**(2), 131.

Bjørgen, Johan P., Sukhorukov, Andrii V., Leenaarts, Jorrit, Carlsson, Mats, de la Cruz Rodríguez, Jaime, Scharmer, Göran B. & Hansteen, Viggo H. (2018), 'Three-dimensional modeling of the ca ii h and k lines in the solar atmosphere', *A&A* **611**, A62.
URL: *https://doi.org/10.1051/0004-6361/201731926*

Boteler, D. H. (2019), 'A 21st century view of the march 1989 magnetic storm', *Space Weather* **17**(10), 1427–1441.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019SW002278*

Cabral, B. & Leedom, L. C. (1993), 'Imaging vector fields using line integral convolution', **1**(1).
URL: *https://www.osti.gov/biblio/10185520*

Camporeale, E. (2019), 'The Challenge of Machine Learning in Space Weather: Nowcasting and Forecasting', *Space Weather* **17**(8), 1166–1207.

Chen, F., Cheung, M. C. M., Rempel, M. & Chintzoglou, G. (2023), 'Data-driven Radiative Magnetohydrodynamics Simulations with the MURaM Code', *The Astrophysical Journal* **949**(2), 118.

Chen, Yajie, Peter, Hardi, Przybylski, Damien, Tian, Hui & Zhang, Jiale (2022), 'Doppler shifts of spectral lines formed in the solar transition region and corona', *A&A* **661**, A94.
URL: *https://doi.org/10.1051/0004-6361/202243111*

Cheung, M. C. M. & Isobe, H. (2014), 'Flux emergence (theory)', *Living Reviews in Solar Physics* **11**(1), 3.
URL: *https://doi.org/10.12942/lrsp-2014-3*

Chian, A. C.-L., Rempel, E. L., Aulanier, G., Schmieder, B., Shadden, S. C., Welsch, B. T. & Yeates, A. R. (2014), 'Detection of coherent structures in photospheric turbulent flows',

The Astrophysical Journal **786**(1), 51.

URL: *http://dx.doi.org/10.1088/0004-637X/786/1/51*

Chian, A. C.-L., Silva, S. S. A., Rempel, E. L., Gošić, M., Bellot-Rubio, L. R., Kusano, K., Miranda, R. A. & Requerey, I. S. (2019), 'Supergranular turbulence in the quiet Sun: Lagrangian coherent structures', *Monthly Notices of the Royal Astronomical Society* **488**(3), 3076–3088.

URL: *https://doi.org/10.1093/mnras/stz1909*

Chian, A. C.-L., Silva, S. S. A., Rempel, E. L., Rubio, L. R. B., Gošić, M., Kusano, K. & Park, S.-H. (2020), 'Lagrangian chaotic saddles and objective vortices in solar plasmas', *Phys. Rev. E* **102**, 060201.

URL: *https://link.aps.org/doi/10.1103/PhysRevE.102.060201*

Cyamahat, W. C. (2022), 'Sun spectral irradiance measured with a calibrated spectrometer (flame s-xr1-es, ocean insight) mounted with a cosine corrector. the measurement was made on a sunny day on 13 june 2022 at noon from hauterive, switzerland. for photometry and colorimetry, standard measurements are usually carried out in the range 360 - 830 nm (highlighted area). correlated color temperature (cct): 5470 k.'. File: Solar spectral irradiance.svg.

URL: *https://en.wikipedia.org/wiki/File:Solar_spectral_irradiance.svg#metadata*

Dakanalis, I., Tsiropoula, G., Tziotziou, K. & Koutroumbas, K. (2021), 'Automated Detection of Chromospheric Swirls Based on Their Morphological Characteristics', *Solar Physics* **296**(1), 17.

De Pontieu, B., Title, A. M., Lemen, J. R., Kushner, G. D., Akin, D. J., Allard, B., Berger, T., Boerner, P., Cheung, M., Chou, C., Drake, J. F., Duncan, D. W., Freeland, S., Heyman, G. F., Hoffman, C., Hurlburt, N. E., Lindgren, R. W., Mathur, D., Rehse, R., Sabolish, D., Seguin, R., Schrijver, C. J., Tarbell, T. D., Wülser, J. P., Wolfson, C. J., Yanari, C., Mudge, J., Nguyen-Phuc, N., Timmons, R., van Bezooijen, R., Weingrod, I., Brookner, R., Butcher, G., Dougherty, B., Eder, J., Knagenhjelm, V., Larsen, S., Mansir, D., Phan, L.,

Boyle, P., Cheimets, P. N., DeLuca, E. E., Golub, L., Gates, R., Hertz, E., McKillop, S., Park, S., Perry, T., Podgorski, W. A., Reeves, K., Saar, S., Testa, P., Tian, H., Weber, M., Dunn, C., Eccles, S., Jaeggli, S. A., Kankelborg, C. C., Mashburn, K., Pust, N., Springer, L., Carvalho, R., Kleint, L., Marmie, J., Mazmanian, E., Pereira, T. M. D., Sawyer, S., Strong, J., Worden, S. P., Carlsson, M., Hansteen, V. H., Leenaarts, J., Wiesmann, M., Aloise, J., Chu, K. C., Bush, R. I., Scherrer, P. H., Brekke, P., Martinez-Sykora, J., Lites, B. W., McIntosh, S. W., Uitenbroek, H., Okamoto, T. J., Gummin, M. A., Auker, G., Jerram, P., Pool, P. & Waltham, N. (2014), 'The Interface Region Imaging Spectrograph (IRIS)', *Solar Physics* **289**(7), 2733–2779.

Denker, C. & Wang, H. (1998), 'High spatial resolution observations of a small δ spot', *The Astrophysical Journal* **502**(1), 493.
URL: *https://dx.doi.org/10.1086/305902*

Fisher, G. H. & Welsch, B. T. (2008), FLCT: A Fast, Efficient Method for Performing Local Correlation Tracking, *in* R. Howe, R. W. Komm, K. S. Balasubramaniam & G. J. D. Petrie, eds, 'Subsurface and Atmospheric Influences on Solar Activity', Vol. 383 of *Astronomical Society of the Pacific Conference Series*, p. 373.

Fleisch, D. (2008), *A Student's Guide to Maxwell's Equations*, Student's Guides, Cambridge University Press.
URL: *https://books.google.co.uk/books?id=I-x1MLny6y8C*

Freytag, B., Steffen, M., Ludwig, H. G., Wedemeyer-Böhm, S., Schaffenberger, W. & Steiner, O. (2012), 'Simulations of stellar convection with CO5BOLD', *Journal of Computational Physics* **231**(3), 919–959.

Ganaie, M. A., Hu, M., Malik, A. K., Tanveer, M. & Suganthan, P. N. (2021), 'Ensemble deep learning: A review', *arXiv e-prints* p. arXiv:2104.02395.

Gary, G. A. (2001), 'Plasma beta above a solar active region: Rethinking the paradigm',

Solar Physics **203**(1), 71–86.

URL: *https://doi.org/10.1023/A:1012722021820*

Gizon, L. & Birch, A. C. (2005), 'Local helioseismology', *Living Reviews in Solar Physics* **2**(1), 6.

URL: *https://doi.org/10.12942/lrsp-2005-6*

Goedbloed, J., Goedbloed, J. & Poedts, S. (2004), *Principles of Magnetohydrodynamics: With Applications to Laboratory and Astrophysical Plasmas*, Cambridge University Press.

URL: *https://books.google.co.uk/books?id=FvM6rMJob-cC*

Graftieaux, L., Michard, M. & Grosjean, N. (2001), 'Combining piv, pod and vortex identification algorithms for the study of unsteady turbulent swirling flows', *Measurement Science and Technology* **12**(9), 1422.

URL: *https://dx.doi.org/10.1088/0957-0233/12/9/307*

Gudiksen, B. V., Carlsson, M., Hansteen, V. H., Hayek, W., Leenaarts, J. & Martínez-Sykora, J. (2011), 'The stellar atmosphere simulation codebifrost', *Astronomy & Astrophysics* **531**, A154.

URL: *http://dx.doi.org/10.1051/0004-6361/201116520*

Hale, G. E., Ellerman, F., Nicholson, S. B. & Joy, A. H. (1919), 'The Magnetic Polarity of Sun-Spots', *The Astrophysical Journal* **49**, 153.

Haller, G. (2015*a*), 'Lagrangian coherent structures', *Annual Review of Fluid Mechanics* **47**(1), 137–162.

URL: *https://doi.org/10.1146/annurev-fluid-010313-141322*

Haller, G. (2015*b*), 'Lagrangian coherent structures', *Annual Review of Fluid Mechanics* **47**(1), 137–162.

URL: *https://doi.org/10.1146/annurev-fluid-010313-141322*

Haller, G., Hadjighasem, A., Farazmand, M. & Huhn, F. (2016), 'Defining coherent vortices objectively from the vorticity', *Journal of Fluid Mechanics* **795**, 136–173.

Haller, G. & Yuan, G. (2000), 'Lagrangian coherent structures and mixing in two-dimensional turbulence', *Physica D: Nonlinear Phenomena* **147**(3), 352–370.
URL: *https://www.sciencedirect.com/science/article/pii/S0167278900001421*

Hildreth, E. (1983), *The Measurement of Visual Motion*, MIT.

Hodges, J. L. (1958), 'The significance probability of the smirnov two-sample test', *Arkiv för Matematik* **3**, 469–486.
URL: *https://api.semanticscholar.org/CorpusID:121451525*

Hornik, K. (1991), 'Approximation capabilities of multilayer feedforward networks', *Neural networks* **4**(2), 251–257.

Hotta, H. & Iijima, H. (2020), 'On rising magnetic flux tube and formation of sunspots in a deep domain', *Monthly Notices of the Royal Astronomical Society* **494**(2), 2523–2537.
URL: *https://doi.org/10.1093/mnras/staa844*

Hotta, H., Iijima, H. & Kusano, K. (2019), 'Weak influence of near-surface layer on solar deep convection zone revealed by comprehensive simulation from base to surface', *Science Advances* **5**(1), eaau2307.
URL: *https://www.science.org/doi/abs/10.1126/sciadv.aau2307*

Iijima, H., Matsumoto, T., Hotta, H. & Imada, S. (2023), 'A comprehensive simulation of solar wind formation from the solar interior: Significant cross-field energy transport by interchange reconnection near the sun', *The Astrophysical Journal Letters* **951**(2), L47.
URL: *http://dx.doi.org/10.3847/2041-8213/acdde0*

Ioffe, S. & Szegedy, C. (2015), 'Batch normalization: Accelerating deep network training by reducing internal covariate shift'.
URL: *https://arxiv.org/abs/1502.03167*

Irbah, A., Borgnino, J., Djafer, D., Damé, L. & Keckhut, P. (2016), 'Solar seeing monitor misolfa: A new method for estimating atmospheric turbulence parameters', *A&A* **591**, A150.
URL: *https://doi.org/10.1051/0004-6361/201527914*

Jeong, J. & Hussain, F. (1995), 'On the identification of a vortex', *Journal of Fluid Mechanics* **285**, 69–94.

Jess, D. B., Morton, R. J., Verth, G., Fedun, V., Grant, S. D. T. & Giagkiozis, I. (2015), 'Multiwavelength Studies of MHD Waves in the Solar Chromosphere. An Overview of Recent Results', *Space Science Reviews* **190**(1-4), 103–161.

Khomenko, E. & Collados, M. (2006), 'Numerical Modeling of Magnetohydrodynamic Wave Propagation and Refraction in Sunspots', *The Astrophysical Journal* **653**(1), 739–755.

Khomenko, E., Collados, M., Díaz, A. & Vitas, N. (2014), 'Fluid description of multicomponent solar partially ionized plasma', *Physics of Plasmas* **21**(9).

Klimchuk, J. A. (2006), 'On solving the coronal heating problem', *Solar Physics* **234**(1), 41–77.
URL: *https://doi.org/10.1007/s11207-006-0055-z*

Künzel, H. (1960), 'Die Flare-Häufigkeit in Fleckengruppen unterschiedlicher Klasse und magnetischer Struktur', *Astronomische Nachrichten* **285**(5), 271.

Lang, K. (2007), *Sun, Earth and Sky*, Springer New York.
URL: *https://books.google.co.uk/books?id=WlHNCRwbQdUC*

Lecun, Y., Bottou, L., Bengio, Y. & Haffner, P. (1998), 'Gradient-based learning applied to document recognition', *Proceedings of the IEEE* **86**(11), 2278–2324.

Lennard, M. G., Silva, S. S. A., Tremblay, B., Ramos, A. A., Verth, G., Ballai, I., Iijima, H., Hotta, H., Rempel, M., Park, S.-H. & Fedun, V. (2025), 'Recovering coherent flow structures in active regions using machine learning', *Monthly Notices of the Royal Astronomical Society* **539**(3), 2498–2512.
URL: *https://doi.org/10.1093/mnras/staf576*

Liu, H., Jin, Z., Xiang, Y. & Ji, K. (2022), 'High-resolution Solar Image Reconstruction Based on Non-rigid Alignment', *Research in Astronomy and Astrophysics* **22**(9), 095005.

Löptien, B., Birch, A. C., Duvall, T. L., Gizon, L., Proxauf, B. & Schou, J. (2017), 'Measuring solar active region inflows with local correlation tracking of granulation', *Astronomy & Astrophysics* **606**, A28.

Löptien, B., Birch, A. C., Duvall, T. L., Gizon, L. & Schou, J. (2016), 'The shrinking sun: A systematic error in local correlation tracking of solar granulation', *Astronomy & Astrophysics* **590**, A130.
URL: *https://doi.org/10.1051/0004-6361/201628112*

Lugt, H. J. (1979), *The Dilemma of Defining a Vortex*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 309–321.
URL: *https://doi.org/10.1007/978-3-642-67220-0_32*

Löfdahl, M. G., Hillberg, T., de la Cruz Rodríguez, J., Vissers, G., Andriienko, O., Scharmer, G. B., Haugan, S. V. H. & Fredvik, T. (2021), 'Sstred: Data- and metadata-processing pipeline for chromis and crisp', *Astronomy & Astrophysics* **653**, A68.
URL: *http://dx.doi.org/10.1051/0004-6361/202141326*

Martínez Pillet, V., del Toro Iniesta, J. C., Álvarez-Herrero, A., Domingo, V., Bonet, J. A., González Fernández, L., López Jiménez, A., Pastor, C., Gasent Blesa, J. L., Mellado, P., Piqueras, J., Aparicio, B., Balaguer, M., Ballesteros, E., Belenguer, T., Bellot Rubio, L. R., Berkefeld, T., Collados, M., Deutsch, W., Feller, A., Girela, F., Grauf, B., Heredero, R. L., Herranz, M., Jerónimo, J. M., Laguna, H., Meller, R., Menéndez, M., Morales, R., Orozco Suárez, D., Ramos, G., Reina, M., Ramos, J. L., Rodríguez, P., Sánchez, A., Uribe-Patarroyo, N., Barthol, P., Gandorfer, A., Knoelker, M., Schmidt, W., Solanki, S. K. & Vargas Domínguez, S. (2011), 'The Imaging Magnetograph eXperiment (IMaX) for the Sunrise Balloon-Borne Solar Observatory', *Solar Physics* **268**(1), 57–102.

Maxwell, J. C. (1865), 'A Dynamical Theory of the Electromagnetic Field', *Philosophical Transactions of the Royal Society of London Series I* **155**, 459–512.

McCulloch, W. S. & Pitts, W. (1943), 'A logical calculus of the ideas immanent in nervous

activity', *The bulletin of mathematical biophysics* **5**(4), 115–133.
URL: *https://doi.org/10.1007/BF02478259*

Minsky, M., Papert, S. & Bottou, L. (2017), *Perceptrons : an introduction to computational geometry*, [2017 edition]. edn, The MIT Press, Cambridge, Massachusetts.

Müller, D., St. Cyr, O. C., Zouganelis, I., Gilbert, H. R., Marsden, R., Nieves-Chinchilla, T., Antonucci, E., Auchère, F., Berghmans, D., Horbury, T. S., Howard, R. A., Krucker, S., Maksimovic, M., Owen, C. J., Rochus, P., Rodriguez-Pacheco, J., Romoli, M., Solanki, S. K., Bruno, R., Carlsson, M., Fludra, A., Harra, L., Hassler, D. M., Livi, S., Louarn, P., Peter, H., Schühle, U., Teriaca, L., del Toro Iniesta, J. C., Wimmer-Schweingruber, R. F., Marsch, E., Velli, M., De Groof, A., Walsh, A. & Williams, D. (2020), 'The Solar Orbiter mission. Science overview', *Astronomy & Astrophysics* **642**, A1.

Nair, V. & Hinton, G. E. (2010), Rectified linear units improve restricted boltzmann machines, *in* 'Proceedings of the 27th International Conference on International Conference on Machine Learning', ICML'10, Omnipress, Madison, WI, USA, p. 807–814.

Newton, H. W. & Nunn, M. L. (1951), 'The Sun's Rotation Derived from Sunspots 1934–1944 and Additional Results', *Monthly Notices of the Royal Astronomical Society* **111**(4), 413–421.
URL: *https://doi.org/10.1093/mnras/111.4.413*

November, L. J. & Simon, G. W. (1988), 'Precise Proper-Motion Measurement of Solar Granulation', *The Astrophysical Journal* **333**, 427.

Osborne, C. M. J. & Milić, I. (2021), 'The Lightweaver Framework for Nonlocal Thermal Equilibrium Radiative Transfer in Python', *The Astrophysical Journal* **917**(1), 14.

Padinhatteeri, S., Sridharan, R. & Sankarasubramanian, K. (2010), 'Seeing-induced errors in solar doppler velocity measurements', *Solar Physics* **266**(1), 195–207.
URL: *http://dx.doi.org/10.1007/s11207-010-9597-1*

Parenti, S. (2014), 'Solar prominences: Observations', *Living Reviews in Solar Physics* **11**(1), 1.

URL: *https://doi.org/10.12942/lrsp-2014-1*

Parker, E. N. (1963), 'Kinematical Hydromagnetic Theory and its Application to the Low Solar Photosphere.', *The Astrophysical Journal* **138**, 552.

Pesnell, W. D., Thompson, B. J. & Chamberlin, P. C. (2012), 'The solar dynamics observatory (sdo)', *Solar Physics* **275**(1), 3–15.

URL: *https://doi.org/10.1007/s11207-011-9841-3*

Potts, H. E., Barrett, R. K. & Diver, D. A. (2004), 'Balltracking: An highly efficient method for tracking flow fields', *A&A* **424**(1), 253–262.

URL: *https://doi.org/10.1051/0004-6361:20035891*

Priest, E. (2014), *Magnetohydrodynamics of the Sun*, Cambridge University Press.

URL: *https://books.google.co.uk/books?id=BrbSAgAAQBAJ*

Przybylski, D., Cameron, R., Solanki, S. K., Rempel, M., Leenaarts, J., Anusha, L. S., Witzke, V. & Shapiro, A. I. (2022), 'Chromospheric extension of the MURaM code', *Astronomy & Astrophysics* **664**, A91.

Raouafi, N. E., Matteini, L., Squire, J., Badman, S. T., Velli, M., Klein, K. G., Chen, C. H. K., Matthaeus, W. H., Szabo, A., Linton, M., Allen, R. C., Szalay, J. R., Bruno, R., Decker, R. B., Akhavan-Tafti, M., Agapitov, O. V., Bale, S. D., Bandyopadhyay, R., Battams, K., Berčič, L., Bourouaine, S., Bowen, T. A., Cattell, C., Chandran, B. D. G., Chhiber, R., Cohen, C. M. S., D'Amicis, R., Giacalone, J., Hess, P., Howard, R. A., Horbury, T. S., Jagarlamudi, V. K., Joyce, C. J., Kasper, J. C., Kinnison, J., Laker, R., Liewer, P., Malaspina, D. M., Mann, I., McComas, D. J., Niembro-Hernandez, T., Nieves-Chinchilla, T., Panasenco, O., Pokorný, P., Pusack, A., Pulupa, M., Perez, J. C., Riley, P., Rouillard, A. P., Shi, C., Stenborg, G., Tenerani, A., Verniero, J. L., Viall, N., Vourlidas,

A., Wood, B. E., Woodham, L. D. & Woolley, T. (2023), 'Parker Solar Probe: Four Years of Discoveries at Solar Cycle Minimum', *Space Science Reviews* **219**(1), 8.

Rempel, E. L., Chertovskih, R., Davletshina, K. R., Silva, S. S. A., Welsch, B. T. & Chian, A. C.-L. (2022), 'Reconstruction of photospheric velocity fields from highly corrupted data', *The Astrophysical Journal* **933**(1), 2.
URL: *https://dx.doi.org/10.3847/1538-4357/ac6fe4*

Rempel, M. (2015), 'Numerical simulations of sunspot decay: On the penumbra–evershed flow–moat flow connection', *The Astrophysical Journal* **814**(2), 125.
URL: *https://dx.doi.org/10.1088/0004-637X/814/2/125*

Rempel, M. & Cheung, M. C. M. (2014), 'Numerical simulations of active region scale flux emergence: From spot formation to decay', *The Astrophysical Journal* **785**(2), 90.
URL: *https://dx.doi.org/10.1088/0004-637X/785/2/90*

Rempel, M., Schüssler, M. & Knölker, M. (2009), 'Radiative Magnetohydrodynamic Simulation of Sunspot Structure', *The Astrophysical Journal* **691**(1), 640–649.

Rieutord, M., Roudier, T., Ludwig, H. G., Nordlund, A. & Stein, R. (2001), 'Are granules good tracers of solar surface velocity fields?', *Astronomy & Astrophysics* **377**, L14–L17.

Rimmele, T. R., Warner, M., Keil, S. L., Goode, P. R., Knölker, M., Kuhn, J. R., Rosner, R. R., McMullin, J. P., Casini, R., Lin, H., Wöger, F., von der Lühe, O., Tritschler, A., Davey, A., de Wijn, A., Elmore, D. F., Fehlmann, A., Harrington, D. M., Jaeggli, S. A., Rast, M. P., Schad, T. A., Schmidt, W., Mathioudakis, M., Mickey, D. L., Anan, T., Beck, C., Marshall, H. K., Jeffers, P. F., Oschmann, J. M., Beard, A., Berst, D. C., Cowan, B. A., Craig, S. C., Cross, E., Cummings, B. K., Donnelly, C., de Vanssay, J.-B., Eigenbrot, A. D., Ferayorni, A., Foster, C., Galapon, C. A., Gedrites, C., Gonzales, K., Goodrich, B. D., Gregory, B. S., Guzman, S. S., Guzzo, S., Hegwer, S., Hubbard, R. P., Hubbard, J. R., Johansson, E. M., Johnson, L. C., Liang, C., Liang, M., McQuillen, I., Mayer, C., Newman, K., Onodera, B., Phelps, L., Puentes, M. M., Richards, C., Rimmele,

L. M., Sekulic, P., Shimko, S. R., Simison, B. E., Smith, B., Starman, E., Sueoka, S. R., Summers, R. T., Szabo, A., Szabo, L., Wampler, S. B., Williams, T. R. & White, C. (2020), 'The daniel k. inouye solar telescope –observatory overview', *Solar Physics* **295**(12), 172.
URL: *https://doi.org/10.1007/s11207-020-01736-7*

Rosenblatt, F. (1958), 'The perceptron: A probabilistic model for information storage and organization in the brain.', *Psychological Review* **65**(6), 386–408.
URL: *https://doi.org/10.1037/h0042519*

Roudier, T., Rieutord, M., Malherbe, J. M. & Vigneau, J. (1999), 'Determination of horizontal velocity fields at the sun's surface with high spatial and temporal resolution', *Astronomy & Astrophysics* **349**, 301–311.

Rouppe van der Voort, L. H. M., De Pontieu, B., Carlsson, M., de la Cruz Rodríguez, J., Bose, S., Chintzoglou, G., Drews, A., Froment, C., Gošić, M., Graham, D. R., Hansteen, V. H., Henriques, V. M. J., Jafarzadeh, S., Joshi, J., Kleint, L., Kohutova, P., Leifsen, T., Martínez-Sykora, J., Nóbrega-Siverio, D., Ortiz, A., Pereira, T. M. D., Popovas, A., Quintero Noda, C., Sainz Dalda, A., Scharmer, G. B., Schmit, D., Scullion, E., Skogsrud, H., Szydlarski, M., Timmons, R., Vissers, G. J. M., Woods, M. M. & Zacharias, P. (2020), 'High-resolution observations of the solar photosphere, chromosphere, and transition region - a database of coordinated iris and sst observations', *Astronomy & Astrophysics* **641**, A146.
URL: *https://doi.org/10.1051/0004-6361/202038732*

Rumelhart, D. E., Hinton, G. E. & Williams, R. J. (1986), 'Learning representations by back-propagating errors', *Nature* **323**, 533–536.
URL: *https://api.semanticscholar.org/CorpusID:205001834*

Scharmer, G. B., Bjelksjo, K., Korhonen, T. K., Lindberg, B. & Petterson, B. (2003), The 1-m Swedish solar telescope, *in* S. L. Keil & S. V. Avakyan, eds, 'Innovative Telescopes and Instrumentation for Solar Astrophysics', Vol. 4853, International Society for Optics and Photonics, SPIE, pp. 341 – 350.
URL: *https://doi.org/10.1117/12.460377*

Scharmer, G., Bjelksjö, K., Korhonen, T., Lindberg, B., Petterson, B., Ab, S., Oy, O. & Ab, L. (2002), 'The 1-meter swedish solar telescope', *Proceedings of SPIE - The International Society for Optical Engineering* .

Scherrer, P. H., Schou, J., Bush, R. I., Kosovichev, A. G., Bogart, R. S., Hoeksema, J. T., Liu, Y., Duvall, T. L., Zhao, J., Title, A. M., Schrijver, C. J., Tarbell, T. D. & Tomczyk, S. (2012), 'The Helioseismic and Magnetic Imager (HMI) Investigation for the Solar Dynamics Observatory (SDO)', *Solar Physics* **275**(1-2), 207–227.

Schmidt, W., von der Lühe, O., Volkmer, R., Denker, C., Solanki, S., Balthasar, H., Gonzalez, N. B., Berkefeld, T., Collados, M., Hofmann, A., Kneer, F., Lagg, A., Puschmann, K., Schmidt, D., Sobotka, M., Soltau, D. & Strassmeier, K. (2012), 'The gregor solar telescope on tenerife'.
URL: *https://arxiv.org/abs/1202.4289*

Schuck, P. W. (2005), 'Local correlation tracking and the magnetic induction equation', *The Astrophysical Journal* **632**(1), L53.
URL: *https://dx.doi.org/10.1086/497633*

Schuck, P. W. (2008), 'Tracking vector magnetograms with the magnetic induction equation', *The Astrophysical Journal* **683**(2), 1134–1152.
URL: *http://dx.doi.org/10.1086/589434*

Schwenn, R. (2006), 'Space weather: The solar perspective', *Living Reviews in Solar Physics* **3**(1), 2.
URL: *https://doi.org/10.12942/lrsp-2006-2*

Shelyag, S., Keys, P., Mathioudakis, M. & Keenan, F. P. (2011), 'Vorticity in the solar photosphere', *Astronomy & Astrophysics* **526**, A5.
URL: *https://doi.org/10.1051/0004-6361/201015645*

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., kin Wong, W. & chun Woo, W. (2015), 'Convo-

lutional lstm network: A machine learning approach for precipitation nowcasting'.
URL: *https://arxiv.org/abs/1506.04214*

Silje, B. (2008), Spicules, Master's thesis, University of Oslo.
URL: *http://urn.nb.no/URN:NBN:no-20856*

Silva, S. S. A., Lennard, M., Verth, G., Ballai, I., Rempel, E. L., Warnecke, J., Iijima, H., Hotta, H., Park, S.-H., Donea, A. C., Kusano, K. & Fedun, V. (2023), 'Novel approach to forecasting photospheric emergence of active regions', *The Astrophysical Journal Letters* **948**(2), L24.
URL: *https://dx.doi.org/10.3847/2041-8213/acd007*

Simon, G. W., Brandt, P. N., November, L. J., Shine, R. A. & Strous, L. H. (1995), Warning: Local Correlation Tracking may BE Dangerous to your (scientific) Health, *in* J. T. Hoeksema, V. Domingo, B. Fleck & B. Battrick, eds, 'Helioseismology', Vol. 376 of *ESA Special Publication*, p. 223.

Singer, H. J., Heckman, G. R. & Hirman, J. W. (2001), 'Space weather forecasting: A grand challenge', *Geophysical Monograph Series* **125**, 23–29.

Sitzmann, V., Martel, J. N. P., Bergman, A. W., Lindell, D. B. & Wetzstein, G. (2020), 'Implicit Neural Representations with Periodic Activation Functions', *arXiv e-prints* p. arXiv:2006.09661.

Stein, R. F. & Nordlund, Å. (1998), 'Simulations of solar granulation. i. general properties', *The Astrophysical Journal* **499**(2), 914.
URL: *https://dx.doi.org/10.1086/305678*

Stein, R. F. & Nordlund, A. k. (2012), 'On the formation of active regions', *The Astrophysical Journal* **753**(1), L13.
URL: *http://dx.doi.org/10.1088/2041-8205/753/1/L13*

Stein, R. F., Nordlund, A. k., Collet, R. & Trampedach, R. (2024), 'The stagger code for accurate and efficient, radiation-coupled magnetohydrodynamic simulations', *The Astro-*

physical Journal **970**(1), 24.

URL: *https://dx.doi.org/10.3847/1538-4357/ad4706*

Strecker, H. & Bello González, N. (2022), 'Evolution of the flow field in decaying active regions - ii. converging flows at the periphery of naked spots', *A&A* **664**, A195.

URL: *https://doi.org/10.1051/0004-6361/202142564*

Strous, L. H. (1995), Feature Tracking: Deriving Horizontal Motion and More, *in* J. T. Hoeksema, V. Domingo, B. Fleck & B. Battrick, eds, 'Helioseismology', Vol. 376 of *ESA Special Publication*, p. 213.

Strous, L. H. & Zwaan, C. (1999), 'Phenomena in an emerging active region. ii. properties of the dynamic small-scale structure', *The Astrophysical Journal* **527**(1), 435.

URL: *https://dx.doi.org/10.1086/308071*

Testa, P., Bakke, H., van der Voort, L. R. & Pontieu, B. D. (2023), 'High-resolution observations of the low atmospheric response to small coronal heating events in an active region core', *The Astrophysical Journal* **956**(2), 85.

URL: *https://dx.doi.org/10.3847/1538-4357/acf4f1*

Toriumi, S., Hayashi, K. & Yokoyama, T. (2012), 'Detection of the horizontal divergent flow prior to the solar flux emergence', *The Astrophysical Journal* **751**(2), 154.

URL: *https://dx.doi.org/10.1088/0004-637X/751/2/154*

Toriumi, S. & Wang, H. (2019), 'Flare-productive active regions', *Living Reviews in Solar Physics* **16**(1), 3.

URL: *https://doi.org/10.1007/s41116-019-0019-7*

Tremblay, B. & Attie, R. (2020), 'Inferring plasma flows at granular and supergranular scales with a new architecture for the deepvel neural network', *Frontiers in Astronomy and Space Sciences* **7**, 25.

URL: *https://www.frontiersin.org/article/10.3389/fspas.2020.00025*

Tremblay, B., Thierry, R., Rieutord, M. & Vincent, A. (2018), 'Reconstruction of horizontal plasma motions at the photosphere from intensitygrams: A comparison between deepvel, lct, flct, and cst', *Solar Physics* **293**, 57.

Tremblay, Benoit, Cossette, Jean-François, Kazachenko, Maria D., Charbonneau, Paul & Vincent, Alain (2021), 'Inferring depth-dependent plasma motions from surface observations using the deepvel neural network', *J. Space Weather Space Clim.* **11**, 9.
URL: *https://doi.org/10.1051/swsc/2020073*

Tsuneta, S., Ichimoto, K., Katsukawa, Y., Nagata, S., Otsubo, M., Shimizu, T., Suematsu, Y., Nakagiri, M., Noguchi, M., Tarbell, T., Title, A., Shine, R., Rosenberg, W., Hoffmann, C., Jurcevich, B., Kushner, G., Levay, M., Lites, B., Elmore, D., Matsushita, T., Kawaguchi, N., Saito, H., Mikami, I., Hill, L. D. & Owens, J. K. (2008), 'The Solar Optical Telescope for the Hinode Mission: An Overview', *Solar Physics* **249**(2), 167–196.

Tziotziou, K., Scullion, E., Shelyag, S., Steiner, O., Khomenko, E., Tsiropoula, G., Canivete Cuissa, J. R., Wedemeyer, S., Kontogiannis, I., Yadav, N., Kitiashvili, I. N., Skirvin, S. J., Dakanalis, I., Kosovichev, A. G. & Fedun, V. (2023), 'Vortex Motions in the Solar Atmosphere', *Space Science Reviews* **219**(1), 1.

Ulrich, R. K. (1970), 'The Five-Minute Oscillations on the Solar Surface', *The Astrophysical Journal* **162**, 993.

van Driel-Gesztelyi, L. & Green, L. M. (2015), 'Evolution of active regions', *Living Reviews in Solar Physics* **12**(1), 1.
URL: *https://doi.org/10.1007/lrsp-2015-1*

Vassiliadis, D., Bothmer, V. & Daglis, I. (2007), *Forecasting space weather*, pp. 403–425.

Verma, M. & Denker, C. (2011), 'Horizontal flow fields observed in hinode g-band images - i. methods', *Astronomy & Astrophysics* **529**, A153.
URL: *https://doi.org/10.1051/0004-6361/201016358*

Verma, M., Steffen, M. & Denker, C. (2013), 'Evaluating local correlation tracking using CO5BOLD simulations of solar granulation', *Astronomy & Astrophysics* **555**, A136.

Vernazza, J. E., Avrett, E. H. & Loeser, R. (1981), 'Structure of the solar chromosphere. III. Models of the EUV brightness components of the quiet sun.', *The Astrophysical Journal Supplement Series* **45**, 635–725.

Vögler, A., Shelyag, S., Schüssler, M., Cattaneo, F., Emonet, T. & Linde, T. (2005), 'Simulations of magneto-convection in the solar photosphere* - equations, methods, and results of the muram code', *A&A* **429**(1), 335–351.
URL: *https://doi.org/10.1051/0004-6361:20041507*

Wachter, R., Schou, J., Rabello-Soares, M. C., Miles, J. W., Duvall, T. L. & Bush, R. I. (2012), 'Image quality of the helioseismic and magnetic imager (hmi) onboard the solar dynamics observatory (sdo)', *Solar Physics* **275**(1), 261–284.
URL: *https://doi.org/10.1007/s11207-011-9709-6*

Ward, F. (1965), 'The General Circulation of the Solar Atmosphere and the Maintenance of the Equatorial Acceleration.', *The Astrophysical Journal* **141**, 534.

Wedemeyer-Böhm, S., Ludwig, H. G., Steffen, M., Freytag, B. & Holweger, H. (2005), 'The shock-patterned solar chromosphere in the light of ALMA'.
URL: *https://arxiv.org/abs/astro-ph/0509747*

Wedemeyer-Böhm, S. & Rouppe van der Voort, L. (2009), 'Small-scale swirl events in the quiet sun chromosphere', *Astronomy & Astrophysics* **507**(1), L9–L12.
URL: *https://doi.org/10.1051/0004-6361/200913380*

Wedemeyer, S., Fleishman, G., de la Cruz Rodríguez, J., Gunár, S., da Silva Santos, J. M., Antolin, P., Guevara Gómez, J. C., Szydlarski, M. & Eklund, H. (2022), 'Prospects and challenges of numerical modeling of the sun at millimeter wavelengths', *Frontiers in Astronomy and Space Sciences* **9**.
URL: *https://doi.org/10.3389/fspas.2022.967878*

Welsch, B. T., Fisher, G. H., Abbett, W. P. & Regnier, S. (2004), 'Ilct: Recovering photospheric velocities from magnetograms by combining the induction equation with local correlation tracking', *The Astrophysical Journal* **610**(2), 1148.
URL: *https://dx.doi.org/10.1086/421767*

Wiegelmann, T. & Sakurai, T. (2012), 'Solar force-free magnetic fields', *Living Reviews in Solar Physics* **9**(1), 5.
URL: *https://doi.org/10.12942/lrsp-2012-5*

Wootten, A. & Thompson, A. R. (2009), 'The Atacama Large Millimeter/Submillimeter Array', *IEEE Proceedings* **97**(8), 1463–1471.

Zhang, S., Do, C.-T., Doddipatla, R., Loweimi, E., Bell, P. & Renals, S. (2021), 'Train your classifier first: Cascade neural networks training from upper layers to lower layers'.
URL: *https://arxiv.org/abs/2102.04697*

Zhou, J., Adrian, R. J., Balachandar, S. & Kendall, T. M. (1999), 'Mechanisms for generating coherent packets of hairpin vortices in channel flow', *Journal of Fluid Mechanics* **387**, 353–396.

Zirker, J. B. (1993), 'Photospheric Vortices and Coronal Heating', *Solar Physics* **147**(1), 47–53.