Probing the solar coronal magnetic field with physics-informed neural networks

Robert Jarolim (robert.jarolim@uni-graz.at)
University of Graz https://orcid.org/0000-0002-9309-2981
Julia Thalmann
University of Graz
Astrid Veronig
University of Graz https://orcid.org/0000-0003-2073-002X
Tatiana Podladchikova
Skolkovo Institute of Science and Technology

Article

Keywords:

Posted Date: April 6th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1415262/v1

Probing the solar coronal magnetic field with physics-informed neural networks

R. Jarolim[1]* J. K. Thalmann[1] A. M. Veronig[1,2] T. Podladchikova[3]

[1]University of Graz, Institute of Physics, Universitätsplatz 5, 8010 Graz, Austria
[2]University of Graz, Kanzelhöhe Observatory for Solar and Environmental Research, Kanzelhöhe 19, 9521 Treffen am Ossiacher See, Austria
[3]Skolkovo Institute of Science and Technology, Bolshoy Boulevard 30, bld. 1, Moscow 121205, Russia

Abstract

While the photospheric magnetic field of our Sun is routinely measured, its extent into the upper atmosphere remains elusive. We present a novel approach for coronal magnetic field extrapolation, using a neural network that integrates observational data and the physical force-free magnetic field model. Our method flexibly finds a trade-off between the observation and force-free magnetic field assumption, improving the understanding of the connection between the observation and the underlying physics. We utilize meta-learning concepts to simulate the evolution of active region NOAA 11158. Our simulation of 5 days of observations at full cadence, requires less than 13 hours of total computation time, enabling real-time force-free magnetic field extrapolations. A systematic comparison of the time evolution of free magnetic energy and magnetic helicity in the coronal volume, as well as comparison to EUV observations demonstrates the validity of our approach. The obtained temporal and spatial depletion of free magnetic energy unambiguously relates to the observed flare activity.

1 Introduction

Extrapolation and simulation methods are used to provide an estimation of the three-dimensional distribution of the coronal magnetic field (1), which is essential to understand the genesis and initiation of solar eruptions and to predict the occurrence of high-energy events from our Sun (2). The most realistic approximations to the magnetized plasma under coronal equilibrium conditions (gas pressure and other forces are negligible), are nonlinear force-free (NLFF) magnetic field models (3). The computation of a force-free magnetic field, B, inside a simulation volume, requires the numerical solution of the system of differential equations

$$J \times B = 0, \tag{1}$$

where J is the electric current density, and

$$\nabla \cdot B = 0 \tag{2}$$

subject to values specified through direct observations on the volume's bounding surface. In practice data at sufficiently high spatial and temporal resolution are obtained from measurements of B at photospheric heights only (at $z = 0$; bottom boundary of the simulation). The data are known to be inconsistent with the force-free assumption (e.g., 4). For compensation, these data are usually specifically processed prior to extrapolation (e.g., 5; 6), and/or the NLFF solution may deviate from the supplied boundary data at $z = 0$ (e.g, 7; 8; 9; 10).

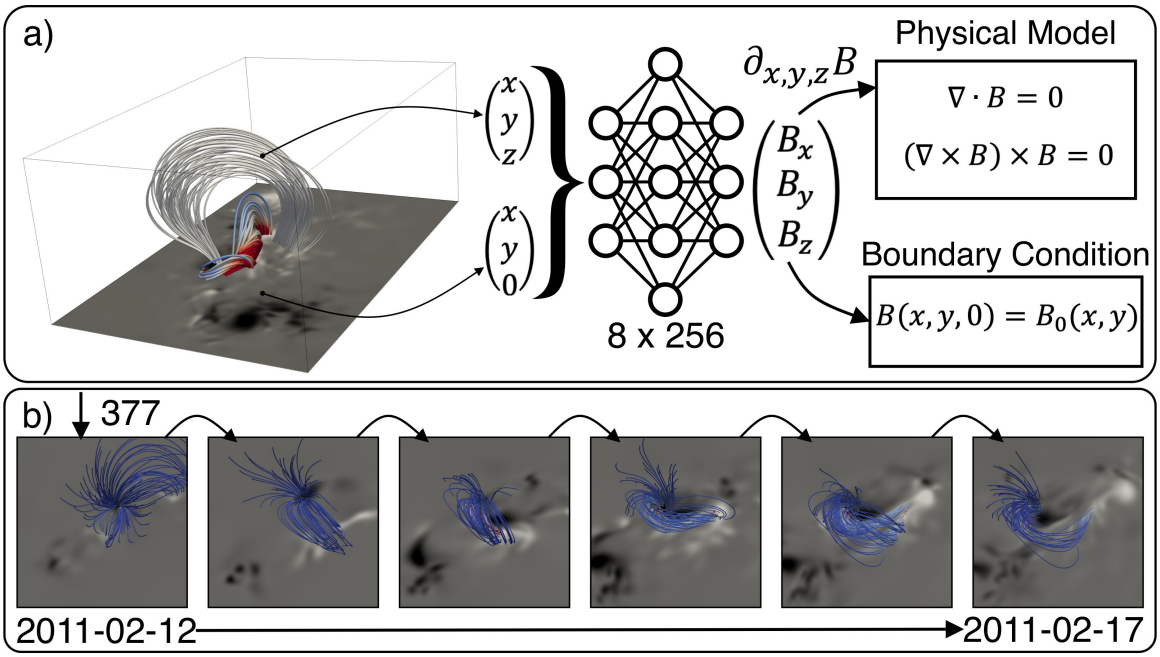


Figure 1: **Overview of the proposed method for force-free magnetic field extrapolation.** a) Our neural network serves as representation of the magnetic field, mapping each coordinate point to the corresponding magnetic field vector (B_x, B_y, B_z). We sample points from the boundary-condition and simulation volume. The neural network is optimized to match observations of the vector magnetic field at the photospheric boundary, while simultaneously minimizing the Lorentz force and divergence in the entire simulation volume, leading to the extrapolation of the magnetic field to higher layers. b) For the simulation of time-series we perform an extrapolation of the first frame from scratch. The consecutive frames use the previous results as starting condition. This allows to efficiently adapt changes and rapidly converge to the next solution. We illustrate the evolution of active region 11158 over five days, and trace the field lines of the negative polarity sunspot. The series shows the evolution of a potential-like field into a strong current carrying flux-rope and its decay back.

The boundary values on the remaining (side and top) boundaries of a usually cubic computational volume are most commonly pre-set by a potential field solution, which can be computed from the observed data at $z = 0$.

Deep learning applications are commonly used to learn a general task by using vast amounts of labeled data (supervised training). Physics-informed neural networks (PINNs; 11) represent a different class that can solve nonlinear partial differential equations through a data-driven approach, uniquely integrating physical models and data in the framework of deep learning. This method has shown the ability to find meaningful solutions even when the problem is not perfectly posed (12), and has been successfully applied for various problem settings (e.g., 13; 14; 15).

We aim at finding a solution to the boundary-value problem by including the NLFF physical equations in a neural network. Here, the neural network acts as function that maps each coordinate point in the simulation volume (x, y, z) to the corresponding magnetic field vector (B_x, B_y, B_z). At the simulation boundary, we match the values of our neural network with the grid-points of the observation $B_0(x, y)$, such that we obtain a neural representation that is a smooth approximation of the boundary magnetic field. For the entire volume, we optimize our model to satisfy the force-free assumption as given by the system of differential equations (Eq.

1 and 2), which determines the extrapolation of the observed magnetogram. The derivative of B with respect to the input coordinates is computed from the neural representation using the automatic differentiation. In addition to the force-free constraint, we require that the magnetic field matches a potential field at the lateral and top boundary (c.f., 7). The model is updated iteratively to minimize the deviation from the boundary-condition and from the physical model (Fig. 1).

The assumption of a force-free magnetic field does not hold for photospheric heights (4; 16) and observations are subject to noise, instrumental characteristics and the stokes inversion that is being used (17). Therefore, both the data and the underlying physical model of our optimization problem are imperfect. Physics informed neural networks have the advantage to enable an intrinsic adjustment of the boundary condition and to allow for deviations from the physical model, where the solution is determined by the weighting between both components (see Methods; Sect. 4).

2 Results

We apply our method to the active region NOAA 11158 (SHARP 377). This region was observed from its emergence (2011-02-12), over more than five days by the Helioseismic and Magnetic Imager (HMI; 18) onboard the Solar Dynamcics Observatory (SDO; 19). The active region developed into a complex quadrupolar sunspot formation that led to multiple solar flare eruptions, including an X2.2-class flare (SOL2011-02-15T01:44; peak-time: 2011-02-15 01:56), originating from the main polarity inversion line along which a sigmoidal filament was observed. An M2.2-class flare (SOL2011-02-14T17:20; peak-time: 2011-02-14 17:26) originated from the eastern part of the active region (for a detailed analysis see (20)).

2.1 Trade-off between observation and physical model

In Fig. 2 we study the trade-off between the physical model and observed data, where we define the weighting parameters for the divergence-free-, force-free-, and boundary-condition, as λ_{div}, λ_{ff}, and λ_B, respectively. In panel b, we compare nine different parameter settings, with varying $\lambda_{div/ff}$, full and factor 2 binned resolution of the input magnetogram, and different conditions for the lateral and top boundary. The plots on the left show the divergence-free (upper) and force-free (lower) error over the difference to the observation data (B_{diff}; Sect. 4). The plots on the right show the normalized metrics, $|\nabla \cdot B|/|B|$ and the average angle between the currents and magnetic field (θ_J; Sect. 4). Therefore for all plots, points towards the left are in better agreement with the observed photospheric magnetic field, while points towards the bottom are in better agreement with the physical model. All metrics show that the ideal point in the lower-left is not attainable for the given boundary-value problem. The performed simulations are distributed along a supposed trade-off line, that determines the attainable solutions. This necessary trade-off is known from existing state-of-the-art NLFF modeling approaches (e.g., 21). Enforcing the constrains of the physical model leads to smoother solutions for the magnetic field with diminished signal on smaller spatial scales, while reducing the physical constrains gives a better approximation of the observation (Fig. 2a).

The use of the full-resolution observational data shows that the higher spatial information only results in a shift along the trade-off line. This suggests that the degree of force-freeness does not improve from the higher spatial resolution.

The usage of a potential field as a boundary condition at the lateral and top boundaries of the model volume, as commonly used in existing NLFF methods, is not strictly required for our method. In Fig. 2b the simulation with an open upper-boundary leads to a supposed improvement orthogonal to the trade-off line (dashed line). We found that for regions of weak-fields this constrain is necessary to prevent artificial fields close to the upper boundary. Therefore, an additional constrain of the lateral and top boundary is required

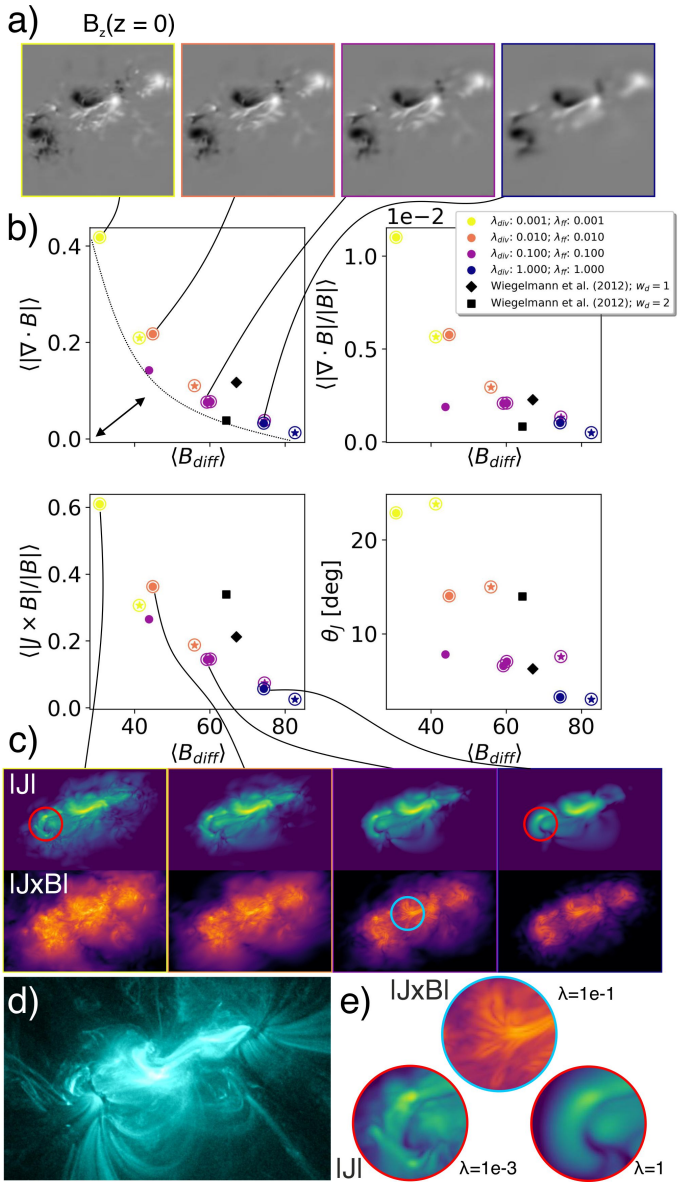


Figure 2: **Evaluation of different parameter settings and trade-off between measurement and physical model**. a) Increasing the weighting of the physical model $\lambda_{diff/ff}$ decreases the similarity to the observed magnetic field. The adjustment of the boundary is intrinsic to satisfy the force-free assumption. b) Metrics on the y-axis represent deviations from the physical model. The x-axis gives the deviation from the observation. Encircled symbols refer to simulations that use a potential-field upper-boundary, stars refer to simulations that use the full spatial resolution, and black symbols refer to a classical state-of-the-art NLFF method for comparison (App. A). The dashed line illustrates the supposed trade-off. c) Maps of integrated current density (top) and integrated horizontal component of $|J \times B|$ (bottom). The different parameter settings result in similar magnetic typologies, where small spatial structures are enhanced when we reduce the constrains of the physical model (bottom of panel e). Deviations from the physical model occur at spatially confined regions and appear highly structured (top of panel e). d) Reference EUV observations of the SDO/AIA 131 Å channel.

to provide valid solutions. The improvement orthogonal to the trade-off line suggests that the potential-field approximation causes a further trade-off with respect to the observed photospheric magnetic field.

Independent of the tested choices of the weighting parameters (λ_{div}, λ_{ff}, and λ_B), all simulations reveal a similar magnetic topology as illustrated by the maps of vertically integrated current density (Fig. 2c). The main difference occurs at smaller scales, where the field appears more structured when the constrains of the physical model are loosened (Fig. 2e). The strong current carrying magnetic fields close to the polarity inversion lines can be also observed in the corresponding EUV observation in Fig. 2d.

In Fig. 2c we show the residual of the horizontal $|J \times B|$ term, which we associate with regions where the force-free assumption can not be satisfied by our model. Relaxing the physical constrains leads to an overall increase of the error, but also to a separation into smaller regions. The blue cutout in Fig. 2e shows the central flux rope, where the largest deviations occur at the intersection of magnetically separated regions (e.g., flux rope and overlaying field).

The optimization allows for a certain variability between individual runs. In Fig. 2b we compare two runs with $\lambda_{div/ff} = 0.1$, which demonstrates that our method converges to similar solutions, and that significant differences only originate from the selected weighting.

2.2 Simulation of time series

In order to obtain a robust understanding of the temporal evolution of the solar corona prior to eruptive events, the computation of time sequences of NLFF models is required. Traditional state-of-the-art methods do so by solving the NLFF boundary value problem individually for each time step. Here, we assume that the model parameters of our neural network are close to temporally adjacent extrapolation results. We start with a full simulation of the initial frame of the series, while the following time-steps are constructed by using the previous simulation as initial model parameters (Fig 1b). This enables a convergence to the next solution that is typically faster than the instrumental cadence (see Sect. 4.3). Figure 3a demonstrates that this training method (solid line) leads to similar results, as performing a training run from scratch (blue symbols).

We apply our method (using $\lambda_{div/ff} = 0.1$) to 601 consecutive magnetograms at a time cadence of 12 minutes, covering the time period from 2011-02-12 00:00 to 2011-02-17 00:00. In Fig. 3, we show the corresponding time profiles of total magnetic energy, free magnetic energy and magnetic helicity (see Sect. 4), together with the GOES soft X-ray flux (bottom) as a reference for the occurrence of large flares. Throughout the evolution of the active region, the total energy gradually increases. In the time profile of the free energy, and even more pronounced in that of the helicity of the current-carrying field, clear responses to the occurrence of solar flares are identified. The free energy density is in agreement with other methods, that report that free-energy is stored low in solar corona (20). The sharp jump in the free energy density (15 02:00) is consistent with a flare related implosion, and can be observed as contraction of the coronal loops (Movie 1).

In Fig. 3b we show spatial maps of depleted free magnetic energy during two flare events, where we integrate along the x-, y-, and z-axis, respectively. Here, we only consider energy losses, which allow for an easier identification of the flares as from the integrated quantities in Fig. 3a. The top row shows the integrated EUV emission over the duration of the flare, which emphasize the regions of energy release (App. 4.1). The right column shows the X-class flare, that occurred close to the central polarity inversion line. The left column shows the M-class flare that originated from the smaller polarity inversion line in the eastern part of the active region. For this event, the flare energy release was not visible in the spatially integrated free energy curves. For both flares the free energy depletion aligns with the position of the flare. We note that also regions revealing a small energy release (light blue), can be associated with enhanced EUV emission (e.g., central flux rope at 14 17:26). For the M-class flare an energy release can be noted at the center, aligning with the minor brightness enhancement in EUV.

The central flux rope becomes clearly visible in AIA 171 Å, minutes before the major flare event (pink circle; Fig. 3c). Similar the magnetic structure related to the null point, indicated by a green circle in Fig.

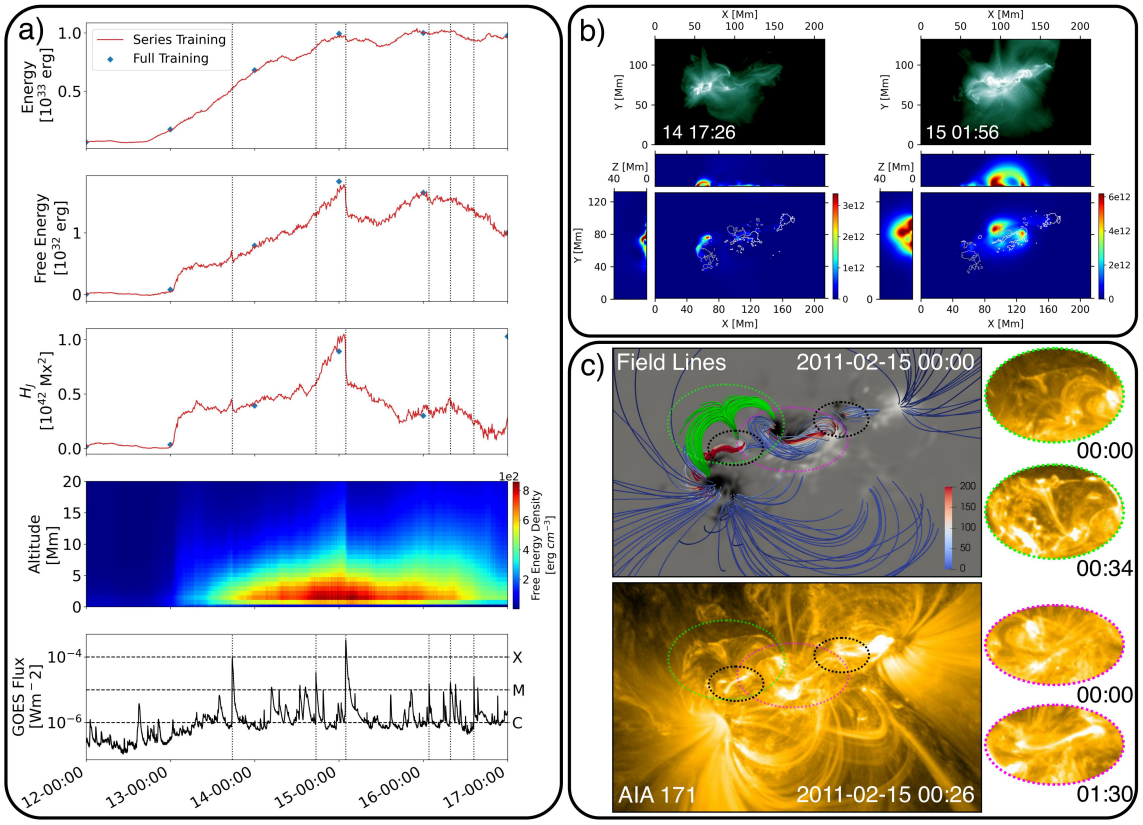


Figure 3: **Simulated series of AR 11158 and comparison of extracted parameters to observations.** a) We show the integrated total magnetic energy, free magnetic energy, magnetic helicity (H_J), free energy density distribution in height, and GOES 1.0–8.0 Å flux. Blue symbols indicate simulations from scratch, that lead to similar results as the series training (red line). The signature of the X-class flare is clearly visible as sharp jump in free energy and helicity. Flares ≥M1 are indicated by vertical lines. The associated contraction of loops can be seen in the height distribution of free energy. b) Integrated EUV observations (top) and the corresponding maps of released free magnetic energy during the flare event (bottom; erg cm^{-2}). Contour lines at ±1000 Gauss are overlayed in gray (negative) and white (positive). The M-class flare is given on the left (2011-02-14 17:26) and the X-class flare on the right (2011-02-15 01:56). The spatial location of free magnetic energy depletion, as estimated by our model, is in agreement with the position of the flare event. c) Magnetic field line traces of the central components of the active region. The colors of the field lines indicate the current density. The two flux ropes associated to pores are encircled in black, the central flux rope in pink, and the eruptive structure associated to the null point in green. A movie of the full series is included in the supplementary material (Movie 1).

3c, appears very faint in the EUV emission and becomes prominent close to the eruption (15 00:34), where we note a good agreement with our modeled magnetic field. Our model suggests that this structure is build up over hours before the eruption. We provide a more extensive comparison in the supplementary material (App. A). In the supplementary movie 1, we show the evolution of the free magnetic energy, current density and running difference of free magnetic energy, for the entire series.

According to our model, the majority of free-energy is stored a few Mm above the polarity inversion line (c.f., supplementary movie 1), where also a large fraction of magnetic energy is depleted during the flare event (Fig. 3b). The change in free energy during the X-class flare above $\sim$ 5Mm is likely related to the contraction of the magnetic field (magnetic implosion; 22). The M-class flare shows an energy release only from heights below below $\sim$ 5Mm. Both events show the same characteristic energy release along the flux rope (c.f., Fig. 3c).

3 Discussion

The comparison between model-deduced flare related changes and coronal (EUV) observations, suggests a suitable approximation of the coronal magnetic field. As evaluated by standard metrics our method achieves state-of-the-art NLFF magnetic field extrapolations (Fig. 2b). The further application to observational data from 66 different solar active regions shows that our method delivers high-quality solutions, independent of the magnetic configuration and area of the considered active region (App. C), and provides the ability to achieve magnetic field extrapolations in quasi real-time (see Sect. 4.3). Extracted parameters from vector magnetograms are frequently used for solar flare prediction (23) and parameters from 3D vector magnetograms have been demonstrated to be reliable precursors for large flares (24). With the fast extrapolation of time series, our method can provide the data sets required to extend these efforts towards data-driven methods, and automatic prediction of solar flares.

A notable advantage is the file size of the resulting extrapolations, which requires about 2 MB. This corresponds to a compression by a factor of 60, as compared to storing the mesh representation, which allows for easy distribution. Furthermore arbitrary regions can be sub-sampled, and smooth derivatives can be computed, by sampling specific coordinates from the neural network.

As shown by the estimated trade-off line (Fig. 2b), the primary restrain of our method is the incompatibility between the underlying physical model and the observation. An extension of the physical model could be achieved by, for instance, accounting for the gradient of the gas pressure (non-force-free magnetic field). This would require further observational constrains to provide physically meaningful results, like traced coronal loops (25) or differential emission measure estimates. Flares are very dynamic processes, and therefore even higher time cadences would be needed than are currently available (12 min). Similar as for the spatial extrapolation, the magnetic field of the intermediate time steps could be interpolated by extending towards a time-dependent model. Another natural next step is the application to magneto-hydrodynamic equations.

4 Method

We employ a physics-informed neural network that uses the force-free assumption to simulate the stratification of the active regions magnetic field, based on the observed photospheric vector magnetogram. Here, the neural network acts as a neural representation of the magnetic field at each point in the spatial simulation volume.

We consider the neural network as function mapping from the coordinate domain to the magnetic field vector ($\mathrm{NN} : \mathbb{R}^3 \mapsto \mathbb{R}^3$; with NN being the neural network). The model training is performed for a single active region, by minimizing the deviation from the observed magnetic field vector at the bottom-boundary, the force-free loss (normalized by B)

$$L_{ff} = \frac{||(\nabla \times B) \times B||^2}{||B||^2 + \epsilon}, \tag{3}$$

and the divergence-free loss

$$L_{div} = |\nabla \cdot B|^2, \tag{4}$$

in the entire volume (c.f., Eq. 1 and 2). Here, B refers to the neural network representation of the magnetic field vector at a given point (i.e., $B = \mathrm{NN}(x, y, z)$) and $\epsilon = 1e-7$ is added for numerical stability. We compute the derivatives of the magnetic field with respect to the input variables ($\frac{\partial B_x}{\partial x}$, $\frac{\partial B_x}{\partial y}$, $\frac{\partial B_x}{\partial z}$, ...) using the automatic computation of the gradient using PyTorch. In contrast to grid based approaches (e.g., finite differences), this allows us to directly compute smooth derivatives at arbitrary positions without sampling the surrounding grid cells.

The driver of the simulation is the boundary condition $B(x, y, 0) = B_0(x, y)$, where B_0 refers to the observation. We optimize for this condition by sampling coordinates from the $z = 0$ layer and minimizing the distance to the corresponding observation from the vector magnetogram. In addition, we account for uncertainties in the measurement by subtracting the error map B_{error} and clipping negative losses (i.e., optimizing only points that exceed the error threshold)

$$L_B = \begin{cases} \|B_{diff,0} - B_{error}\|^2, & \text{if } B_{diff,0} - B_{error} > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

Here, $B_{diff,0}$ refers to the difference between the neural network prediction and the observation $|B - B_0|$. In this formulation the lateral and top boundaries of the cubic simulation domain are not explicitly bound. We assume that the magnetic field approaches a potential field solution at those boundaries. For this we compute the potential field solution using the Green's function approach as proposed by (26). This condition is imposed via the same loss terms as for the bottom-boundary (Eq. 5), where we assume no error.

We use Spaceweather HMI Active Region Patches (SHARPs) that are remapped to an cylindrical equal area (CEA; i.e., equidistant coordinate system). The three vector components (B_p, $-B_t$, B_r) serve as bottom-boundary condition. Their corresponding error estimates are used in the loss function as B_{errror}. Per default we bin the data by a factor of two in both spatial dimensions. We normalize the magnetic field strength to 2500 Gauss and the spatial coordinate dimensions to 160 pixels (320 for the original resolution). Values > 1 are considered as input and output. For all our evaluations we transform the magnetic extrapolation back to Gauss units and use the mesh representation by sampling all coordinate points in the simulation volume.

For observations of NOAA 11158 (SHARP 377), we crop the active region patches to x [66:658] and y [9:377] pixels (prior to the binning), in correspondence with the optimization-based (3, see Appendix A) NLFF modeling used for comparison (for a similar model time series see, e.g., 27).

For our model, we adapt the SIREN model architecture from (28), where we use sin activation functions throughout our network. The sin activation showed a slightly better performance than the tanh activation functions, that were originally proposed for PINNs (11). Our model takes the 3 spatial coordinates as input and maps them over 8 fully-connected layers, with 256 channels each, to the 3 output components of the magnetic field vector. We omit the activation function for the output. Our evaluation suggests that a larger number of free parameters does not lead to a performance increase of our method.

While a positional encoding in terms of Fourier features (29) leads to a faster fitting of the high-frequency features of the $z = 0$ vector magnetogram, we found that it is not beneficial for the extrapolation of the magnetic field that is solely determined by the physical equations.

The training is performed iteratively, where we randomly sample 30,000 points from the simulation volume. 10,000 points are taken from grid-positions of the boundary coordinates and are used to compute L_{B0} (Eq. 5). The remaining 20,000 points are randomly and mesh-free sampled from the spatial volume ($x = [0, x_{max}], y = [0, y_{max}], z = [0, 159]$) according to a uniform distribution. The L_{ff} and L_{div} loss is computed from all sampled points (Eq. 3 and 4). We optimize the neural network for the combined loss

$$L = \lambda_{ff} L_{ff} + \lambda_{div} L_{div} + \lambda_{B0} L_{B0}, \tag{6}$$

where the λ parameters refer to the weighting of the individual losses.

The use of the original spatial resolution is straightforward, by using the full resolution of the boundary grid and increasing the sampling height to $z = [0, 319]$. The sampling strategy remains unchanged, resulting in an equal convergence time.

For our model training we use the Adam optimizer with beta of (0.9, 0.999) and apply gradient clipping of 0.1. We train our model for 100,000 iterations until we reach convergence. We found that prioritizing the boundary condition at the beginning of the training leads to a faster convergence. We set $\lambda_{B0} = 1000$ and exponentially decay it to 1 over the first $50,000$ iterations. Similarly, we exponentially decay the learning rate from $5e-4$ to $5e-5$ over the first $70,000$ iterations. The choice of λ_{ff} and λ_{div} is discussed in the main text.

In order to employ the metrics used for quantifying the quality of our NLFF modeling, we sample the full simulation volume and compute the derivatives using finite-differences. From this we compute $\nabla \cdot B$, and $\nabla \times B$.

As an indicator of the degree of force-freeness of the obtained solution, the current-weighted average of the angle between the current density and the magnetic field is used (30), where

$$\theta_J = \sin^{-1} \frac{\sum_i J_i \sigma_i}{\sum_i J_i}. \tag{7}$$

and

$$\sigma_i = \frac{||J \times B||_i}{||J||_i ||B||_i}, \tag{8}$$

and where we employ the corresponding average $\langle \theta_J \rangle$ over all grid points.

We denote the difference, including the error uncertainty B_{error}, at each point r within the observed data B_0 as

$$B_{diff}(r) = |B(r) - B_0(r)| - B_{error}(r). \tag{9}$$

For computing $\langle B_{diff} \rangle$, we first set negative values to 0 and compute the vector norm, such that we estimate the deviation from a tolerance range. Averaging over the full simulation volume then gives $\langle B_{diff} \rangle$.

The total magnetic energy is computed by integrating over the full simulation volume (V)

$$E = \int_V \frac{B^2}{8\pi} dV, \tag{10}$$

where we denote E_{FF} when derived from our force-free solution and E_{PF} when derived from the potential field solution. The free magnetic energy is estimated from the difference between the force-free and potential energy

$$E_{free} = E_{FF} - E_{PF}. \tag{11}$$

For integrated maps of free energy, we apply the Green‘s function method with the bottom magnetic field of our extrapolation to obtain the potential magnetic field. The potential field for the evaluation of the free magnetic energy evolution is obtained by the computationally faster approach based on Fourier expansion by (31).

We note that all metrics can be calculated using the automatic differentiation of the neural representation at each coordinate point, which leads to similar results.

4.1 Integrated Maps

As an observational reference we use EUV filtergrams from the Atmospheric-Imagining-Assembly (AIA; 18) on-board SDO. For the integrated EUV emission in Fig. 3b, we use observations from flare start to end at 1 minute cadence. We use EUV observations from the AIA 94 Å channel, sensitive to the hot flare plasma

(peak formation temperature T = 7 MK), reproject them to the corresponding SHARP and normalize for exposure time. From the set of images we compute the average per pixel, where we mask saturated pixels (>3,000 DN/s). The images are plotted using a logarithmic normalization over the value range [10, 1000]. For the maps of free magnetic energy release, we use the NLFF solutions temporally closest to the nominal GOES start and end times of the flare. We compute the difference per grid cell, set positive differences to zero and integrate along the individual spatial axis. The energies are normalized linearly between 0 and the maximum value of the map that is obtained from integrating along the z-axis (x-y projection).

4.2 Magnetic Helicity

Given that the boundary-normal components of B and B_p match on the surface of the simulation, the gauge-invariant relative magnetic helicity can be written as (32; 33)

$$H_V = \int_V \left(A + A_p\right) \cdot \left(B - B_p\right) dV, \tag{12}$$

where A and A_p are the respective vector potentials satisfying $B = \nabla \times A$ and $B_p = \nabla \times A_p$, as well as $\nabla \cdot A = 0$ and $\nabla \cdot A_p = 0$, respectively. H_V in Eq. (12) can be decomposed as, $H_V = H_J + H_{PJ}$ (34; 35), where

$$H_J = \int_V \left(A - A_p\right) \cdot \left(B - B_p\right) dV, \tag{13}$$

is the magnetic helicity of the current-carrying field, $B_J = B - B_p$, and

$$H_{PJ} = 2 \int_V A_p \cdot \left(B - B_p\right) dV, \tag{14}$$

is the volume-threading helicity. The vector potentials A and A_p are computed using the Coulomb-gauge method of (36).

4.3 Series Training

A primary aspect of force-free magnetic field extrapolations is the application to time series. Existing methods perform the simulation from scratch (e.g., 21; 20). Here, we make use of the relatively small changes in time, by iteratively using the weights of the previous time step as initial weights for training. We find that the neural network rapidly adapts to the changes of the photospheric magnetic field, such that we obtain a result comparable to a simulation from scratch (Fig. 3), in a fraction of the time. For the simulation of a time-series we start by training a neural network from scratch for the first magnetogram of the series. The resulting weights are used as starting point for the subsequent extrapolation (analogously to transfer learning), where we set $\lambda_B = 1$ and train with the new boundary-magnetic field for 2,000 iterations.

For our experiments we used four NVIDIA V100 GPUs. The initial simulation is performed in about 50 minutes (about 2 hours on a single V100 GPU). The consecutive simulations require approximately 1.2 minutes per extrapolation (including data pre-processing and calculating the potential field boundary). This result already allows for operational magnetic-field extrapolations in real-time, even with less computing power. A strategy could be to perform the initial extrapolations with the first sight of active regions on the east limb. Afterwards, the proceeding time-steps can be computed at a fraction of the instrumental cadence. A limitation is the initial training time, that delays the extrapolation results. The use of better initial weights, for example by using meta-learning, could be investigated to further speed-up the initial training time.

5 Code availability

Our codes are publicly available. We provide Python notebooks that perform simulations for arbitrary regions without any pre-requirements. We provide GPU accelerated code for computing the potential field solution based on the Green's function method (26). `https://github.com/RobertJaro/NF2`

6 Data availability

All our simulation results are publicly available (parameter variation, time series, 66 individual active regions). See project page: `https://github.com/RobertJaro/NF2`

The SDO HMI and AIA data is provided by JSOC. We provide automatic download scripts with SunPy (37; 38).

References

[1] Wiegelmann, T., Petrie, G. J. D. & Riley, P. Coronal Magnetic Field Models. **210**, 249–274 (2017).

[2] Green, L. M., Török, T., Vršnak, B., Manchester, W. & Veronig, A. The Origin, Early Evolution and Predictability of Solar Eruptions. **214**, 46 (2018). `1801.04608`.

[3] Wiegelmann, T. & Sakurai, T. Solar force-free magnetic fields. *Living Reviews in Solar Physics* **18**, 1 (2021). `1208.4693`.

[4] Metcalf, T. R., Jiao, L., McClymont, A. N., Canfield, R. C. & Uitenbroek, H. Is the Solar Chromospheric Magnetic Field Force-free? **439**, 474 (1995).

[5] Wiegelmann, T., Inhester, B. & Sakurai, T. Preprocessing of Vector Magnetograph Data for a Nonlinear Force-Free Magnetic Field Reconstruction. **233**, 215–232 (2006). `astro-ph/0612641`.

[6] Fuhrmann, M., Seehafer, N., Valori, G. & Wiegelmann, T. A comparison of preprocessing methods for solar force-free magnetic field extrapolation. **526**, A70 (2011). `1010.6015`.

[7] Wiegelmann, T. & Inhester, B. How to deal with measurement errors and lacking data in nonlinear force-free coronal magnetic field modelling? **516**, A107 (2010).

[8] Wheatland, M. S. & Régnier, S. A Self-Consistent Nonlinear Force-Free Solution for a Solar Active Region Magnetic Field. **700**, L88–L91 (2009). `0906.4414`.

[9] Wheatland, M. S. & Leka, K. D. Achieving Self-consistent Nonlinear Force-free Modeling of Solar Active Regions. **728**, 112 (2011). `1012.3503`.

[10] Wiegelmann, T. *et al.* How Should One Optimize Nonlinear Force-Free Coronal Magnetic Field Extrapolations from SDO/HMI Vector Magnetograms? **281**, 37–51 (2012). `1202.3601`.

[11] Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* **378**, 686–707 (2019).

[12] Karniadakis, G. E. *et al.* Physics-informed machine learning. *Nature Reviews Physics* **3**, 422–440 (2021).

[13] Raissi, M., Yazdani, A. & Karniadakis, G. E. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science* **367**, 1026–1030 (2020).

[14] Mishra, S. & Molinaro, R. Physics informed neural networks for simulating radiative transfer. **270**, 107705 (2021). 2009.13291.

[15] Mathews, A., Hughes, J., Francisquez, M., Hatch, D. & White, A. Uncovering edge plasma dynamics via deep learning of partial observations. In *APS Division of Plasma Physics Meeting Abstracts*, vol. 2020 of *APS Meeting Abstracts*, TO10.007 (2020).

[16] Peter, H., Warnecke, J., Chitta, L. P. & Cameron, R. H. Limitations of force-free magnetic field extrapolations: Revisiting basic assumptions. **584**, A68 (2015). 1510.04642.

[17] Borrero, J. M. *et al.* VFISV: Very Fast Inversion of the Stokes Vector for the Helioseismic and Magnetic Imager. **273**, 267–293 (2011). 0901.2702.

[18] Lemen, J. R. *et al.* The Atmospheric Imaging Assembly (AIA) on the Solar Dynamics Observatory (SDO). **275**, 17–40 (2012).

[19] Pesnell, W. D., Thompson, B. J. & Chamberlin, P. C. The Solar Dynamics Observatory (SDO). *Solar Physics* **275**, 3–15 (2012).

[20] Sun, X. *et al.* Evolution of Magnetic Field and Energy in a Major Eruptive Active Region Based on SDO/HMI Observation. **748**, 77 (2012). 1201.3404.

[21] Thalmann, J. K., Sun, X., Moraitis, K. & Gupta, M. Deducing the reliability of relative helicities from nonlinear force-free coronal models. **643**, A153 (2020). 2009.05287.

[22] Hudson, H. S. Implosions in Coronal Transients. **531**, L75–L77 (2000).

[23] Camporeale, E. The Challenge of Machine Learning in Space Weather: Nowcasting and Forecasting. *Space Weather* **17**, 1166–1207 (2019). 1903.05192.

[24] Kusano, K., Iju, T., Bamba, Y. & Inoue, S. A physics-based method that can predict imminent large solar flares. *Science* **369**, 587–591 (2020).

[25] Wiegelmann, T. & Neukirch, T. Including stereoscopic information in the reconstruction of coronal magnetic fields. **208**, 233–251 (2002). 0801.3234.

[26] Sakurai, T. Green's function methods for potential magnetic fields. *Solar Physics* **76**, 301–321 (1982).

[27] Thalmann, J. K., Linan, L., Pariat, E. & Valori, G. On the Reliability of Magnetic Energy and Helicity Computations Based on Nonlinear Force-free Coronal Magnetic Field Models. **880**, L6 (2019). 1907.01179.

[28] Sitzmann, V., Martel, J., Bergman, A., Lindell, D. & Wetzstein, G. Implicit neural representations with periodic activation functions. *Advances in Neural Information Processing Systems* **33**, 7462–7473 (2020).

[29] Tancik, M. *et al.* Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in Neural Information Processing Systems* **33**, 7537–7547 (2020).

[30] Wheatland, M. S., Sturrock, P. A. & Roumeliotis, G. An Optimization Approach to Reconstructing Force-free Fields. **540**, 1150–1155 (2000).

[31] Alissandrakis, C. On the computation of constant alpha force-free magnetic field. *Astronomy and Astrophysics* **100**, 197–200 (1981).

[32] Berger, M. A. & Field, G. B. The topological properties of magnetic helicity. *Journal of Fluid Mechanics* **147**, 133–148 (1984).

[33] Finn, J. & Antonsen, T. J. Magnetic helicity: What is it and what is it good for? *Comments Plasma Phys. Controlled Fusion* **9**, 111 (1984).

[34] Berger, M. A. Introduction to magnetic helicity. *Plasma Physics and Controlled Fusion* **41**, B167–B175 (1999).

[35] Berger, M. A. *Topological quantities in magnetohydrodynamics*, 345–374 (2003).

[36] Thalmann, J. K., Inhester, B. & Wiegelmann, T. Estimating the Relative Helicity of Coronal Magnetic Fields. **272**, 243–255 (2011).

[37] Mumford, S. J. *et al.* Sunpy (2020). URL `https://doi.org/10.5281/zenodo.3871057`.

[38] Glogowski, K., Bobra, M. G., Choudhary, N., Amezcua, A. B. & Mumford, S. J. drms: A python package for accessing hmi and aia data. *Journal of Open Source Software* **4**, 1614 (2019). URL `https://doi.org/10.21105/joss.01614`.

[39] Zacharov, I. *et al.* "Zhores" — Petaflops supercomputer for data-driven modeling, machine learning and artificial intelligence installed in Skolkovo Institute of Science and Technology. *Open Engineering* **9**, 59 (2019). `1902.07490`.

[40] Barnes, W. T. *et al.* The sunpy project: Open source development and status of the version 1.0 core package. *The Astrophysical Journal* **890**, 68 (2020).

[41] Astropy Collaboration *et al.* Astropy: A community Python package for astronomy. **558**, A33 (2013). `1307.6212`.

[42] Paszke, A. *et al.* Pytorch: An imperative style, high-performance deep learning library. In Wallach, H. *et al.* (eds.) *Advances in Neural Information Processing Systems 32*, 8024–8035 (Curran Associates, Inc., 2019). URL `http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf`.

[43] Ahrens, J., Geveci, B. & Law, C. Paraview: An end-user tool for large data visualization. *The visualization handbook* **717** (2005).

[44] Valori, G., Démoulin, P., Pariat, E. & Masson, S. Accuracy of magnetic energy computations. **553**, A38 (2013). `1303.6773`.

[45] Valori, G. *et al.* Magnetic Helicity Estimations in Models and Observations of the Solar Magnetic Field. Part I: Finite Volume Methods. **201**, 147–200 (2016). `1610.02193`.

7 Acknowledgements

This research has received financial support from the European Union's Horizon 2020 research and innovation program under grant agreement No. 824135 (SOLARNET). JT and AV acknowledge the Austrian Science Fund (FWF): P31413-N27. The authors acknowledge the use of Zhores for obtaining the results presented in this paper (39). This research has made use of SunPy v3.0.0 (37; 40), AstroPy (41), PyTorch (42) and Paraview (43).

8 Author contributions

R.J. developed the method and led the writing of the paper, J.T. performed the evaluation and comparison to existing NLFF methods, and contributed to the writing of the paper, A.V. contributed to the conceptualization of the study and writing of the paper, T.P. contributed to the HPC computations. All authors discussed the results and commented on the manuscript.

A Comparison to other NLFF extrapolations

We compare the performance of our newly proposed method to that of the well-tested and widely used optimization-based extrapolation method of (10). The optimization-based modeling has been carried out for a reduced time cadence of 1-hour, but keeping the native 12-min time cadence of SDO/HMI measurements around two large flares (an M6.6-class flare which peaked at 2011-02-13 17:28 UT and an X2.2-class flare that peaked at 2011-02-15 01:56 UT). The optimization-based modeling has been carried out twofold, differing by the relative weighting of the volume-integrated Lorentz force and divergence, both to be minimized simultaneously based on the optimization principle (parameters w_f and w_d, respectively, in Eq. (4) of 10). Once, an equally strong weighting is chosen ($w_f = w_d = 1$) and once the volume-integrated divergence is weighted stronger ($w_f = 1$, $w_d = 2$).

We consider the $\lambda_{div/ff} = 0.1$ and $\lambda_{div/ff} = 1$ setting, which both show only small deviations from the physical model (Fig. 2b). The $\lambda_{div/ff} = 0.1$ extrapolation has a higher weighting of the boundary-condition.

For analysis of method-based differences, on the one hand we analyze the time evolution of retrieved physical parameters, including the free magnetic energy ($E_{\rm free}$; cf. Sect. 4) and magnetic helicity of the current-carrying field (H_J; cf. Sect. 4.2). On the other hand, we inspect the model quality, by evaluating the degree of force-freeness using $\langle\theta_J\rangle$ (cf. Sec. 4), as well as measures quantifying the deviation of the NLFF solutions from a solenoidal state ($\nabla \cdot B = 0$). Those measures are based on the decomposition of the magnetic energy into solenoidal and non-solenoidal components (44), with $E_{\rm div}/E$ quantifying the fraction of the total magnetic energy related to the non-zero divergence of B. Dedicated studies showed that at most a value of $\simeq 0.1$ is to be tolerated (45; 27).

In Supplementary Fig. 1 we compare the derived parameters and quality metrics, where we note three major differences between the methods.

1. The free magnetic energies of our extrapolations are lower throughout the entire series (this also applies to the total and potential energies). Since all methods have a relatively small divergence, the total energy is mostly determined by the bottom-boundary field.

2. The trend of the helicity deviates between the two methods at the end of the series. Here, our method shows a decrease after the X-class flare, while the method from (10) shows an increasing trend.

3. The time profiles of magnetic energy and helicity of the current-carrying field derived from the application of the method by (10), exhibit a sharp peak at the time of fast magnetic flux (filament) emergence early on Feb 13. Our method shows a more gradual increase until both methods reach similar values in free energy and helicity.

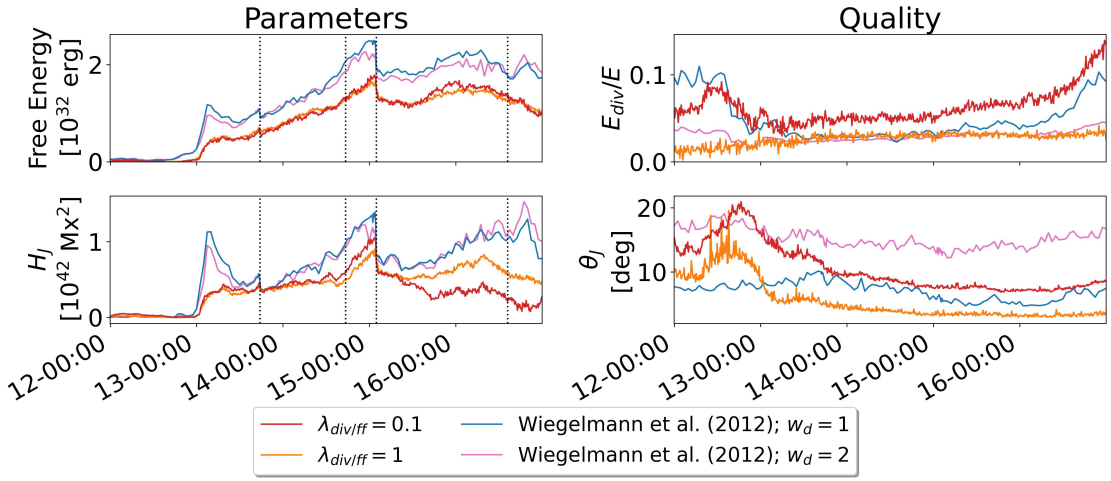
At the same time, we note consistency between the individual method-based results.

1. Both methods show signatures of the major flare events.

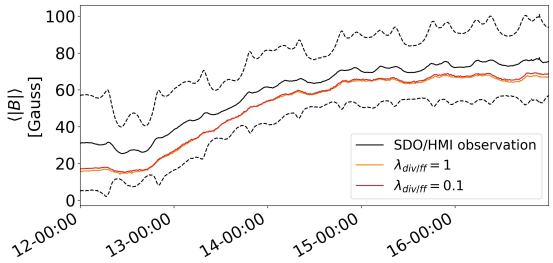
2. Despite the mentioned differences, the trends of free magnetic energy and helicity are similar.

3. In terms of quality metrics the methods are in the same range. $E_{div}/E < 0.1$ is mostly satisfied for all extrapolations. θ_J is initially large for our extrapolations and decreases later in the series. This is associated to the increasing magnetic field strength. The drawback of the better quality metrics (i.e., $\lambda_{div/ff} = 1$) is the larger difference to the observation, which also leads to smoother transitions during the large flare events.

In terms of the differences, current observational capabilities are insufficient to determine which solution is in better agreement with the actual magnetic field.



Supplementary Figure 1: **Comparison of our method to the optimization method of (10).** We evaluate derived physical parameters (left) and quality metrics (right) for the full series. A lower quality score is associated with a better performance.



Supplementary Figure 2: **Comparison of the average magnetic field strength at** $z = 0$.

B Deviation from observed magnetic field strength

The force-free assumption requires a deviation from the observed magnetic field. In Supplementary Fig. 2 we show the temporal evolution of the absolute magnetic field strength at $z = 0$, for both our method and the observed magnetogram. The error range is computed based on the provided error maps, and is given by the dashed lines. For the entire series, our method stays within the uncertainty range.

C Application to different active regions

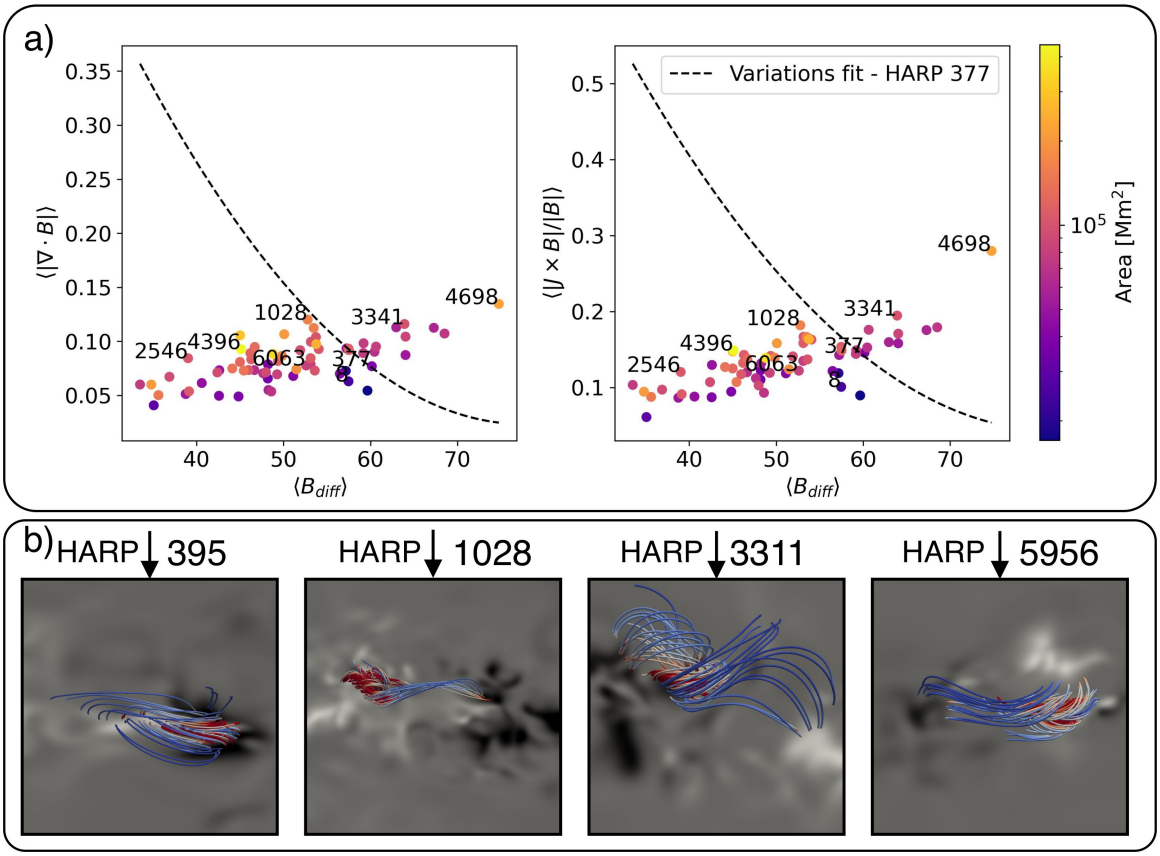
We apply our method to a statistical set of active region patches that are one hour prior to the eruption of an X- or M-class flare. We select only magnetograms that are located within $|x| < 500$ arcsec (helioprojective coordinates), and that provide a valid quality flag. Therefore, we focus on complex active regions, that are the primary subject of magnetic field extrapolations.

We train our model with $\lambda_{div/ff} = 0.1$ and reduce the training time to 80,000 iterations. We found that

we usually reach convergence after 70,000 iterations, where only minor improvements are achieved from continuing the training.

Supplementary Figure 3 shows the evaluation of the results, where we use the metrics from Fig. 2b. As reference we plot a second order polynomial fit of the results from our λ variations, where we consider all extrapolations with the potential field boundary condition. The simulations show a similar performance, and no clear dependence on the area of the active region. This suggests that our method provides stable results for arbitrary active regions. We note that despite our evaluation, extrapolations should be always evaluated for their respective performance and validity. In the supplementary materials we provide a CSV-file summarizing the full set of performance metrics.

As can be clearly seen from Supplementary Fig. 3, the performance scatters orthogonal to the trade-off line, where we would associate simulations towards the bottom-left with a better performance. We associate this trend with the (in)compatibility of the observation with the force-free assumption (data dependent). Therefore, for observations that are closer to a force-free field, the trade-off line shifts towards the bottom-left. The scatter parallel to the trade-off line can be associated with the quality variations of the individual extrapolations (method dependent).



Supplementary Figure 3: **Quality evaluation of our extrapolation method applied to 66 active regions.** a) We plot the error of the divergence-free (left) and force-free (right) assumption over the deviation from the bottom-boundary magnetic field. The dashed line corresponds to the fitting curve from our parameter variations. Selected active regions are labeled across the performance range, including SHARP 377 (NOAA 11158) from our previous evaluation. The color indicates the area of the individual active regions. Extrapolation runs scatter orthogonal to the trade-off line depending on the observed magnetic field. b) The identified flux-ropes of four active regions are illustrated. The color of the field-lines indicates the current density , where red corresponds to strong current carrying fields.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- movie1.mp4