

Optimization of Photospheric Electric Field Estimates for Accurate Retrieval of Total Magnetic Energy Injection

E. Lumme¹  · J. Pomoell¹  · E.K.J. Kilpua¹ 

Received: 29 June 2017 / Accepted: 19 November 2017 / Published online: 7 December 2017
© Springer Science+Business Media B.V., part of Springer Nature 2017

Abstract Estimates of the photospheric magnetic, electric, and plasma velocity fields are essential for studying the dynamics of the solar atmosphere, for example through the derivative quantities of Poynting and relative helicity flux and using the fields to obtain the lower boundary condition for data-driven coronal simulations. In this paper we study the performance of a data processing and electric field inversion approach that requires only high-resolution and high-cadence line-of-sight or vector magnetograms, which we obtain from the *Helioseismic and Magnetic Imager* (HMI) onboard *Solar Dynamics Observatory* (SDO). The approach does not require any photospheric velocity estimates, and the lacking velocity information is compensated for using *ad hoc* assumptions. We show that the free parameters of these assumptions can be optimized to reproduce the time evolution of the total magnetic energy injection through the photosphere in NOAA AR 11158, when compared to recent state-of-the-art estimates for this active region. However, we find that the relative magnetic helicity injection is reproduced poorly, reaching at best a modest underestimation. We also discuss the effect of some of the data processing details on the results, including the masking of the noise-dominated pixels and the tracking method of the active region, neither of which has received much attention in the literature so far. In most cases the effect of these details is small, but when the optimization of the free parameters of the *ad hoc* assumptions is considered, a consistent use of the noise mask is required. The results found in this paper imply that the data processing and electric field inversion approach that uses only the photospheric magnetic field information offers a flexible and straightforward way to obtain photospheric magnetic and electric field estimates suitable for practical applications such as coronal modeling studies.

Keywords Corona, active · Corona, models · Helicity, magnetic · Helicity, observations · Magnetic fields, photosphere · Magnetic fields, corona

✉ E. Lumme
erkka.lumme@helsinki.fi

¹ Department of Physics, University of Helsinki, P.O. Box 64, 00014 Helsinki, Finland

1. Introduction

Determination of the photospheric magnetic, electric, and plasma velocity fields is central for understanding many of the dynamical processes in the solar atmosphere. These fields are required for quantifying the transport of magnetic flux, energy, and helicity from the solar interior through the photosphere to the chromosphere and corona (Liu and Schuck, 2012; Tziotziou, Georgoulis, and Liu, 2013; Kazachenko *et al.*, 2015). They also form the essential input for data-driven modeling of the corona (Wiegmann, Petrie, and Riley, 2015; Inoue, 2016; Leake, Linton, and Schuck, 2017), in which the fields are used to specify the boundary condition at the lower radial boundary of the simulations. Owing to the sparsity of direct observations of many of the key quantities that characterize the coronal plasma – most prominently the magnetic field – data-driven simulations offer currently the most feasible and often the only way to quantify the dynamics of the corona. Consequently, data-driven modeling is the key for studying a variety of coronal phenomena, including coronal mass ejections (CMEs; *e.g.* Chen, 2011; Pagano, Mackay, and Poedts, 2013) and flares (*e.g.* Shibata and Magara, 2011; Cheung and DeRosa, 2012; Fisher *et al.*, 2015; Jiang *et al.*, 2016), and the formation of the solar wind (*e.g.* Riley *et al.*, 2011).

Remote-sensing measurements of the photospheric magnetic field vector and the line-of-sight (LOS) component of the plasma velocity are routinely made at high resolution by both space-based (Tsuneta *et al.*, 2008; Scherrer *et al.*, 2012) and ground-based instruments (*e.g.* Mickey *et al.*, 1996; Keller, Harvey, and Giampapa, 2003). However, the available data do not directly provide the photospheric plasma velocity vector and/or the electric field, and as a result, it is necessary to employ rather complex inversion methods to fully retrieve these fields. A variety of methods exists for this task (Welsch *et al.*, 2007; Ravindra, Longcope, and Abbett, 2008; Schuck, 2008; Kazachenko, Fisher, and Welsch, 2014; Tremblay and Vincent, 2015), all sharing similar basic building blocks: The photospheric velocity/electric fields are constrained to be consistent with the available time series of the magnetic field measurements and the (ideal) magnetohydrodynamic induction equation (or Faraday's law). Moreover, most of the methods employ optical flow methods (*e.g.* Schuck, 2006) to acquire a proxy for the transverse plasma velocity component, for which no direct remote-sensing observations exist. Some of the methods additionally employ the LOS velocity estimates from Dopplergrams (*i.e.* full-disk photospheric Doppler images) as an additional constraint.

In the face of the incomplete input data that inherently include noise as well as calibration issues (Welsch, Fisher, and Sun, 2013; Lagg *et al.*, 2015; Schuck *et al.*, 2016), it is clear that any given inversion method can only yield results up to a certain degree of accuracy. So far, the performance of the methods has been evaluated using synthetic photospheric data extracted from a magnetohydrodynamic (MHD) simulation in which the true plasma velocity and electric fields are known (Abbett *et al.*, 2004; Welsch *et al.*, 2007). The best-performing methods in these tests include the PDFI method of Kazachenko, Fisher, and Welsch (2014) and DAVE4VM by Schuck (2008). In addition to quantifying the degree to which the velocity and electric fields of the simulation are reproduced, the tests of the methods have also focused on the correct reproduction of the fluxes of magnetic energy (*i.e.* Poynting flux) and relative magnetic helicity through the photosphere, as both of these quantities are central for coronal dynamics. Sufficient amount of magnetic energy, and more specifically, free magnetic energy (*i.e.* energy in excess of that of the potential magnetic field configuration), is pivotal for solar eruptions (Aly, 1991; Chen, 2011; Shibata and Magara, 2011). Relative magnetic helicity, that is, the helicity in excess of the helicity of the potential magnetic field configuration, may have a significant role at least in some solar eruptions, although its significance is still under debate (Démoulin, 2007). For example, Tziotziou,

Georgoulis, and Raouafi (2012) used non-linear force-free field modeling for a large number of active regions and found that eruptive active regions tend to accumulate large budgets of both free magnetic energy and signed relative helicity. On the other hand, Pariat *et al.* (2017) recently found using a set of coronal MHD simulations in idealized settings that a large signed relative helicity budget does not provide a reliable eruptivity proxy. Instead they found that the ratio between the relative helicity in the current-carrying part of the field and the total relative helicity best differentiates eruptive from non-eruptive coronal configurations.

The accurate estimation of the Poynting and helicity fluxes is of significant practical value when photospheric velocity/electric fields are used to derive the photospheric boundary condition for data-driven coronal modeling, in which case the boundary condition controls the injection of these quantities to the simulation domain. Such data-driven models vary in complexity, ranging from simple static (non-linear) force-free field models employing only the magnetic field as the boundary condition (see *e.g.* Wiegelmann and Sakurai, 2012) to full time-dependent MHD models employing fully consistent photospheric boundary conditions (*e.g.* Jiang *et al.*, 2016, Leake, Linton, and Schuck, 2017). In terms of complexity, a model exists between these extremes, which is called the time-dependent magnetofrictional method (which we henceforth abbreviate as the TMF method). The TMF method approximates the evolution of the coronal magnetic field as an interplay between two competing processes: relaxation toward a force-free state, and a constantly evolving photospheric boundary condition (Yang, Sturrock, and Antiochos, 1986; van Ballegooijen, Priest, and Mackay, 2000; Cheung and DeRosa, 2012; Pomoell, Lumme, and Kilpua, 2017). Owing to the simplified coronal physics, the method offers a computationally inexpensive way for coronal modeling, but still has sufficient accuracy for time-dependent simulations of the formation and possibly also of the ejection of CME flux ropes (Mackay and van Ballegooijen, 2006; Gibb *et al.*, 2014; Fisher *et al.*, 2015; Weinzierl *et al.*, 2016). The TMF method employs the photospheric electric field as its only data-driven boundary condition, thereby offering a direct application for photospheric electric field estimates. In addition, coronal TMF simulations offer an additional way for evaluating the performance of the inversion methods through studying the realism of the simulation feedback to the electric field boundary condition (*e.g.* Cheung and DeRosa, 2012; Cheung *et al.*, 2015).

In this paper we study how the fluxes of magnetic energy and relative helicity depend on the choice of the electric field inversion method using a data processing and electric field inversion approach that requires only high-resolution and high-cadence (vector) magnetograms as an input and does not use any photospheric velocity estimates. The approach is flexible to use, but the accuracy of the electric field estimates is restricted as the missing photospheric velocity information must be completed using *ad hoc* assumptions. We show that despite these limitations, we are able to optimize the free parameters of the applied electric field inversion methods to yield accurate estimates for the total (and implicitly also the free) magnetic energy injected through the photosphere for a single active region, NOAA AR 11158. This region has been studied extensively by several authors (Schrijver *et al.*, 2011; Cheung and DeRosa, 2012; Jing *et al.*, 2012; Liu and Schuck, 2012; Tziotziou, Georgoulis, and Liu, 2013; Kazachenko *et al.*, 2015). The paper is organized as follows: In Section 2 we present the data processing and electric field inversion methods used in the study and describe how we apply these methods to NOAA AR 11158. In Section 3 we present the main findings of our study and also discuss some of the data processing details that may affect the results, but which have not garnered much attention in the previous works. These include the masking threshold used in removing the noise-dominated pixels and the method of tracking the active region as it traverses across the solar disk. In Section 4

we discuss the implications of the results on the general applicability of our data processing and electric field inversion methods, particularly from the point of view of providing the photospheric boundary condition for coronal (TMF) simulations.

2. Data and Methods

In this section we first describe the methods we used to create time series of photospheric magnetic field and electric field maps, and how we estimate the magnetic energy and relative helicity injection through the photosphere from these maps. We then present the application of these methods to NOAA AR 11158, and describe how we optimized our electric field estimates for an accurate retrieval of the total magnetic energy injection in this active region.

2.1. Data Analysis and Computation Tools

In this study we employed two tools to invert the photospheric electric field, ELECTRICIT and DAVE4VM. ELECTRIC field Inversion Toolkit (ELECTRICIT) is our newly developed software toolkit that handles all steps from downloading and processing of magnetogram data to the inversion of the photospheric electric field (following the data processing methods of Sun, 2013; Welsch, Fisher, and Sun, 2013; Hoeksema *et al.*, 2014, and Kazachenko *et al.*, 2015, explained in Section 2.2, and the electric field inversion methods of Cheung and DeRosa, 2012; Cheung *et al.*, 2015, and Kazachenko, Fisher, and Welsch, 2014, explained in Section 2.3). This paper employs the entire pipeline of the toolkit. The toolkit is written in Python, and its functionality relies heavily on SunPy (Mumford *et al.*, 2015) and SciPy libraries (Jones, Oliphant, and Peterson, 2001) as well as on the FISHPACK code (Swarztrauber and Sweet, 1975), which we call from Python. Currently, the toolkit is limited to electric field inversion based on vector or LOS magnetograms alone.

The differential affine velocity estimator for vector magnetograms (DAVE4VM) is a velocity inversion code developed by Schuck (2008). When fed with a time series of photospheric vector magnetograms, the toolkit determines the photospheric velocity vector field by minimizing the vertical component of the ideal induction equation,

$$\frac{\partial B_z}{\partial t} + \nabla_h \cdot (B_z \mathbf{V}_h - V_z \mathbf{B}_h) = 0, \quad (1)$$

in the neighborhood of each magnetogram pixel. Here z refers to the direction perpendicular to the photosphere, and the subscript h corresponds to the components parallel to the photosphere in a local Cartesian system where the curvature of the solar surface is neglected. The output velocity field can be used to determine the electric field through the ideal Ohm's law:

$$\mathbf{E} = -\mathbf{V} \times \mathbf{B}. \quad (2)$$

In this paper we apply DAVE4VM using vector magnetograms processed by ELECTRICIT (see Section 2.3 for details).

2.2. Processing of Vector Magnetograms

Our photospheric data input consists of full-disk disambiguated vector magnetograms provided by the *Helioseismic and Magnetic Imager* (HMI) instrument (Scherrer *et al.*, 2012)

onboard the *Solar Dynamics Observatory* (SDO) spacecraft (Pesnell, Thompson, and Chamberlin, 2012) described in detail by Hoeksema *et al.* (2014). The magnetograms are provided with a spatial resolution of $0.5''$ per pixel in the plane of sky and a cadence of 720 seconds. In each pixel the magnetic field is inverted from the full Stokes vector observations made at a cadence of 135 seconds averaged over a tapered time window of 1350 seconds. We use ELECTRICIT to download these magnetograms from the Joint Science Operations Center (JSOC) under the dataset title `hmi.B_720s` using a modified version of the download functionality of the SunPy library.

As explained by Hoeksema *et al.* (2014), the vector magnetogram data product includes four different options for the disambiguation of the magnetic field vector. Strong-field pixels (nominally pixels where $B > 150 \text{ Mx cm}^{-2}$) are disambiguated using the minimum energy method (Metcalf, 1994), whereas for weak-field pixels, the user can decide between three disambiguation methods: the potential field acute angle method, the radial acute angle method, or a randomly chosen disambiguation. We chose to use the potential field acute angle method for the weak-field pixels. We recognize the fact that the method can produce unphysical artifacts (Liu *et al.*, 2017). However, the approach constrains the azimuth using a physically justifiable condition (unlike the randomly chosen disambiguation), and it is applicable over the entire solar disk (unlike the radial acute angle method, which in practice yields a random disambiguation near the disk center, see *e.g.* Gosain and Pevtsov, 2013). Since we mask the noise-dominated weak-field pixels to zero before the inversion of the electric field, the choice of the weak-field disambiguation method is not significant for the results presented in this paper (see further discussion in Section 3.3.2).

The HMI vector magnetograms exhibit spurious bad pixels in regions where the Stokes inversion code fails. Since bad pixels can be recognized by exceptionally high formal errors given by the Stokes inversion module (Hoeksema *et al.*, 2014), we used a fixed threshold for the error of the total field strength $\sigma_B = 750 \text{ Mx cm}^{-2}$, above which we labeled each pixel as bad. This threshold recognized all bad pixels in the strong-field region for the test cases we considered (see *e.g.* Hoeksema *et al.*, 2014), but seems to be slightly oversensitive in the weak-field region. We substituted the vector components of the bad pixels (in the HMI image basis, Sun, 2013) by the median of the good pixels in the nearest symmetric neighborhood of each bad pixel that contained at least one good pixel. Bad pixels were searched for and fixed only in a rectangular full-disk cutout that contained the pixels required in interpolating the final reprojected magnetogram (explained below). Appendix C lists further details about the number of bad pixels we removed when creating the reprojected magnetogram times series for NOAA AR 11158 (Section 2.5).

After processing the full-disk magnetograms, ELECTRICIT tracks the user-specified region and creates a time series of reprojected vector magnetograms centered at the given region and projected to local Cartesian coordinates, similarly to the Space-weather HMI Active Region Patches (SHARPs) (Bobra *et al.*, 2014; Hoeksema *et al.*, 2014). The active regions are tracked similarly to SHARP data (Sun, 2013; Hoeksema *et al.*, 2014) with a few modifications discussed further in Appendix A.1. As an additional processing step, we removed the spurious temporal flips in the azimuth of the magnetic field using a slightly modified version of the method presented by Welsch, Fisher, and Sun (2013) (see Appendix A.2 for details).

Finally, before using the time series of processed and reprojected magnetic field vector data in the electric field inversion, we masked the noise-dominated weak-field pixels to zero in order to avoid introducing noise to the solution. Since our electric field inversion methods use three subsequent magnetograms at $t - \Delta t$, t , and $t + \Delta t$ to compute the electric field at time t (Section 2.3), we demand the total magnetic field strength to be higher than the noise

threshold in all three frames. We employed the noise threshold of $B = 250 \text{ Mx cm}^{-2}$ from Kazachenko *et al.* (2015), which is 2.5 times the nominal noise level $\sigma_B = 100 \text{ Mx cm}^{-2}$ in HMI vector magnetograms (Hoeksema *et al.*, 2014). We reproduced the determination of the noise level in the local Cartesian magnetic field components from Kazachenko *et al.* (2015) for our data (see Section 3.3), and because we used different data processing and disambiguation methods, our noise levels are lower. However, as we show in Section 3.3, the noise threshold of the mask has a notable effect on the results, and hence, to enable optimal compatibility to the results of Kazachenko *et al.* (2015), we used the same noise threshold. For DAVE4VM input, we did not mask the data since DAVE4VM cannot handle input data that have large regions of uniform (zero) values. Instead, we masked the data after the velocity inversion was performed.

2.3. Electric Field Inversion

Currently, the electric field inversion within ELECTRICIT is based on a combination of the PTD part from the PTD-Doppler-FLCT-Ideal (PDFI) method (Kazachenko, Fisher, and Welsch, 2014) and *ad hoc* assumptions proposed by Cheung and DeRosa (2012) and Cheung *et al.* (2015). The PDFI approach has yielded good results in tests using synthetic data, and it is also flexible and allows an inversion when only part of the required input data is available (as in our case). At the core of the method is the decomposition of the electric field into inductive \mathbf{E}_I and non-inductive $-\nabla\psi$ components:

$$\mathbf{E} = \mathbf{E}_I - \nabla\psi. \quad (3)$$

The inductive component is completely constrained by a times series of vector magnetograms (from which $\partial\mathbf{B}/\partial t$ can be estimated) and Faraday's law:

$$\nabla \times \mathbf{E}_I = -\frac{\partial\mathbf{B}}{\partial t}, \quad (4)$$

which in the PDFI method is uncurled using a poloidal–toroidal decomposition (PTD). We solved the inductive component similarly to Kazachenko, Fisher, and Welsch (2014) with some modifications to the applied numerical methods (see Appendix B). Our electric field estimates are exactly consistent with Faraday's law and the evolution of B_z in our time series of magnetograms.

The inversion of the non-inductive component is more challenging. In PDFI method, this is done using the ideal Ohm's law (Equation 2) and three constraints from measurements: Dopplergram velocities near polarity inversion lines (PILs), optical flow velocities obtained using the FLCT method (Fisher and Welsch, 2008), and the ideal constraint $\mathbf{E} \cdot \mathbf{B} = 0$ (giving the Doppler-FLCT-Ideal part of the method title). Although this method provides state-of-the-art inversion results in tests, we used a more simplified approach because velocity data processing including the calibration of Dopplergrams and determination of FLCT velocity is yet to be implemented in the ELECTRICIT pipeline. Instead, we used one of the following three (*ad hoc*) assumptions to determine the non-inductive potential ψ :

0. $\nabla\psi = 0$
1. $\nabla_h^2\psi = -\nabla_h \cdot \mathbf{E}_h = -\Omega B_z$
2. $\nabla_h^2\psi = -U j_z = -U(\nabla \times \mathbf{B}) \cdot \hat{\mathbf{z}}$,

where Ω and U are freely chosen constants. The zeroth assumption simply sets the non-inductive component to zero. Fisher *et al.* (2010) studied the performance of such an inversion in reproducing the electric field from simulated photospheric data, and found that the

assumption reproduces the known electric field poorly, *i.e.* the contribution from $-\nabla\psi$ is significant and cannot be approximated to zero. Moreover, Kazachenko, Fisher, and Welsch (2014) estimated that only 40% of the total vertical Poynting flux is reproduced properly when the assumption is used. Nevertheless, we retained this assumption in this study as a point of reference and because there are simulation studies that have reported realistic coronal evolution despite the use of the assumption in deriving the photospheric boundary condition. For example, Gibb *et al.* (2014) modeled a formation of a flux rope in the flaring AR 10977, although this resulted in an only partial agreement with the observed X-ray sigmoid. Thus it is useful to quantify how large the deviations in the magnetic energy and relative helicity fluxes are when this assumption is used in inverting the electric field from real magnetograms. Assumption 1 was conceived by Cheung and DeRosa (2012), who found that Assumption 0 did not produce enough free magnetic energy for their data-driven TMF simulation of flux rope ejections. Assumption 1 was introduced as an *ad hoc* method to increase the free energy budget to more realistic values without making the electric field inversion too complicated. As noted by the authors, the assumption imposes uniform vortical motions to the photospheric plasma. Using the ideal Ohm’s law, we can indeed show that Assumption 1 is obtained for the case of a vertical axisymmetric flux tube rotating at a constant angular velocity $\omega = -\Omega/2$. Assumption 2 was presented by Cheung *et al.* (2015) for their TMF simulation study of recurrent homologous helical jets. They proposed that the jets were driven by an emergence of a twisted current-carrying magnetic structure, and consequently, Assumption 2 is obtained for the case of a uniform emergence ($\mathbf{V} = U\hat{\mathbf{z}}$) of an axisymmetric flux tube, the axis of which is parallel to the z direction, and which is invariant along its axis. In ELECTRICIT we solve the Poisson equations of Assumptions 1 and 2 using the same numerical machinery as in solving the inductive component (see Appendix B for details). The determination of the optimal values for the free parameters Ω and U is discussed in Section 2.6.

In the DAVE4VM-based inversion, we calculate the electric field from the output velocity field of DAVE4VM using the ideal Ohm’s law (Equation 2). For each time t , the input to DAVE4VM consists of a vector magnetogram at time t , estimates for the horizontal derivatives of the vector components at that time (estimated using the five-point optimized derivatives as in Schuck, 2008), and an estimate for $\partial B_z/\partial t$ calculated from the magnetograms at time $t + \Delta t$ and $t - \Delta t$ using a central difference scheme. Defining the input in this way gives the output velocity field estimate also at time t . The window size in DAVE4VM is 19×19 pixels, the same as used by Liu and Schuck (2012), and it maximizes the Pearson correlation between the two terms of the induction equation for the vector magnetogram data we use in this study (Section 2.5). The correlation is not very good (~ 0.6) because of the high noise levels in the input vector magnetograms (Y. Liu, 2017, private communication), and thus DAVE4VM-based electric fields fulfill Faraday’s law poorly. Consequent discrepancies introduced to the time evolution of B_z prevent using these electric fields directly as the data-driven boundary condition of coronal simulations (as discussed by Schuck, 2008).

2.4. Calculation of Magnetic Energy and Relative Helicity Injections Through the Photosphere

The magnetic energy injected through the photosphere E_m to the upper solar atmosphere can be calculated by integrating the vertical Poynting flux S_z over each magnetogram and integrating this total energy flux in time (*e.g.* Kazachenko *et al.*, 2015):

$$E_m(t) = \int_0^t dt' \frac{dE_m}{dt} = \int_0^t dt' \int dA S_z = \frac{1}{\mu_0} \int_0^t dt' \int dA (\mathbf{E} \times \mathbf{B}) \cdot \hat{\mathbf{z}}. \tag{5}$$

The relative helicity H_R injected through the photosphere can be calculated similarly (Berger and Field, 1984; Démoulin, 2007):

$$H_R(t) = \int_0^t dt' \frac{dH_R}{dt} = -2 \int_0^t dt' \int dA (\mathbf{A}_p \times \mathbf{E}) \cdot \hat{\mathbf{z}} \quad (6)$$

$$= -2 \int_0^t dt' \int dA (A_p^x E_y - A_p^y E_x). \quad (7)$$

Here A_p is the vector potential that corresponds to the potential field extrapolation derived from the B_z component. Similarly as in Kazachenko *et al.* (2015), we calculate A_p in ELECTRICIT using the PTD formalism for the (masked) magnetic field using exactly the same numerical machinery as in solving the inductive electric field (Appendix B).

2.5. Implementation of the Analysis Pipeline for NOAA AR 11158

In this paper we employ the analysis pipeline described in Sections 2.1–2.4 to the active region NOAA 11158. The active region emerged on February 10, 2011, in the southern hemisphere of the Sun, and its strongest manifestation of activity was the X2.2 flare (onset on February 15 at 01:44 UT) and the associated halo CME (Schrijver *et al.*, 2011). The magnetic energy and helicity budgets of this active region have been studied extensively using photospheric estimates (Liu and Schuck, 2012; Tziotziou, Georgoulis, and Liu, 2013; Kazachenko *et al.*, 2015) and force-free extrapolations (*e.g.* Jing *et al.*, 2012; Sun *et al.*, 2012; Aschwanden, Sun, and Liu, 2014). The results of both approaches have been reviewed by Kazachenko *et al.* (2015). Cheung and DeRosa (2012) also performed a series of illustrative data-driven simulations for the active region using a *semi-ad hoc* electric field inversion from LOS magnetic field data as the photospheric boundary condition. They employed Assumption 1 from Section 2.3 and tested two values for the free Ω parameter (discussed further in Section 4).

For this study we created our own time series of reprojected vector magnetograms for NOAA AR 11158 using the data processing functionality of ELECTRICIT as described in Section 2.2. The time series started on February 10 00:00 UT well before the emergence of the active region when the active region was at Stonyhurst longitude E53 and lasted until February 20 00:00 UT when the center of the active region was approximately at Stonyhurst longitude W77. The region was tracked at a fixed heliographic latitude S21.0769°. The reprojected vector magnetograms in the series are centered at the active region, and they are projected to the Mercator projection with a projection pixel size of $0.03^\circ \times 0.03^\circ$ that matches the spatial resolution of HMI full-disk vector magnetograms at the disk center. We scaled the vertical component of B_z by $\cos^2 \lambda'$, where λ' is the latitude of the pixel in heliographic coordinates where the center of the reprojected magnetogram patch is at the equator, in order to account for the distortion of the pixel scale in the Mercator projection (Kazachenko *et al.*, 2015). The size of the reprojected magnetograms is 744×610 pixels, which is large enough to include the entire active region. An example magnetogram from the time series is shown in Figure 1.

For the energy and helicity injection studies, we inverted the photospheric electric field for each frame from February 10 14:00 UT until February 18 00:00 UT. We created three electric field time series using ELECTRICIT and each of the *ad hoc* assumptions 0–2 and one time series from the DAVE4VM velocity inversion using the ideal Ohm's law (see Section 2.3 for details). We excluded the data starting from February 18, as from this date

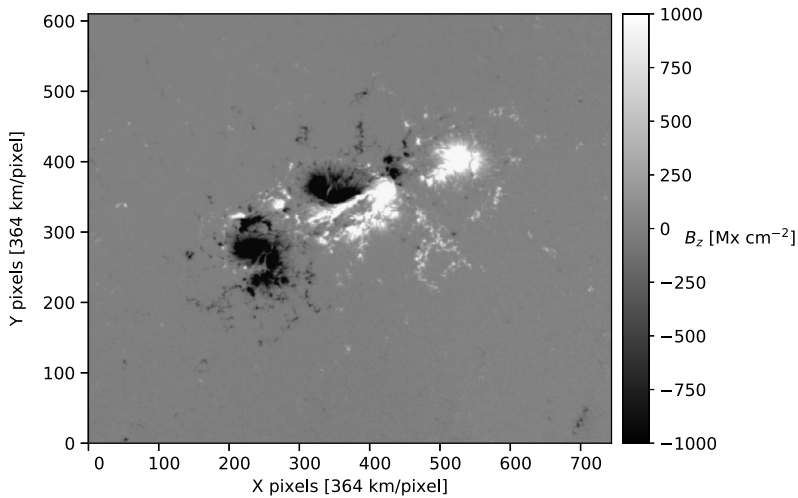


Figure 1 Snapshot from the time series of reprojected vector magnetograms used in this paper. The plot illustrates the B_z component of the magnetic field on February 15 01:36 UT when the center of the active region was at (S21, W13).

onward, the active region was already so far at the limb that the degrading temporal stability of the azimuth disambiguation introduces spurious signals to the electric fields inverted using ELECTRICIT (similar to the “halos” detected by Weinzierl *et al.*, 2016).

2.6. Optimization of the Free Parameters of the Electric Field Inversion

We optimized the free parameters of our electric field inversion (Ω or U , Section 2.3) so that the optimal values reproduced the time evolution of the total magnetic energy injected through the photosphere $E_m(t)$ (Section 2.4, Equation 5) as provided by the reference estimate (explained below) as best as possible. The optimal values of Ω and U are constant throughout the time series of the electric field maps. There is no reason why they could not be time dependent, but considering this option is beyond the scope of this paper. We focus on the optimization of the magnetic energy, because, as discussed in the introduction, the (free) magnetic energy injected through the photosphere is central for the initiation of solar eruptions. We considered the total magnetic energy instead of the free part of the energy for simplicity and because, in our case, the optimization of the total magnetic energy injection in fact includes the optimization of the free energy. It can be shown that modifying the non-inductive component does not affect the injection of magnetic energy to the potential field $E_m^{\text{pot}}(t)$, and thus, by optimizing the total magnetic energy through the non-inductive component $E_m(t) = E_m^{\text{pot}}(t) + E_m^{\text{free}}(t)$ we simultaneously optimize the free-energy part $E_m^{\text{free}}(t)$. Intuitively, this argument seems obvious because the potential magnetic field configuration and the energy within it are determined solely by the photospheric distribution of B_z , which is in no way affected by the choice of $-\nabla\psi$. This also applies when the electric field time series is used to drive a TMF simulation, since the inductive component E_1 (independent of the choice of $-\nabla\psi$) determines the evolution of the photospheric B_z , and thus also the evolution of the potential field. However, this result can be also shown more formally using the decomposition of the vertical Poynting flux into potential and free energy contributions

(Welsch, 2006):

$$S_z = S_z^{\text{pot}} + S_z^{\text{free}} = \frac{1}{\mu_0} [\mathbf{E} \times \mathbf{B}^{\text{pot}} + \mathbf{E} \times (\mathbf{B} - \mathbf{B}^{\text{pot}})] \cdot \hat{\mathbf{z}}, \tag{8}$$

where \mathbf{B}^{pot} corresponds to the magnetic field of the potential field in the photosphere:

$$\mathbf{B}^{\text{pot}} = -\nabla\chi|_{z=0}. \tag{9}$$

Here χ is the scalar potential that fulfills the Laplace equation $\nabla^2\chi = 0$ in the corona with the Neumann boundary condition $\partial\chi/\partial z|_{z=0} = -B_z$ at the photosphere. Welsch (2006) showed that the flux of potential-field energy S_z^{pot} in the decomposition above gives the photospheric energy injection rate to the potential part of the coronal magnetic field dE_m^{pot}/dt . Using similar arguments, it is straightforward to show that dE_m^{pot}/dt is independent of the choice of the non-inductive component:

$$\frac{dE_m^{\text{pot}}}{dt} = \int dA S_z^{\text{pot}} = \frac{1}{\mu_0} \int dA (\mathbf{E} \times \mathbf{B}^{\text{pot}}) \cdot \hat{\mathbf{z}} \tag{10}$$

$$= \frac{1}{\mu_0} \int dA [(\mathbf{E}_I - \nabla\psi) \times (-\nabla\chi)] \cdot \hat{\mathbf{z}} \tag{11}$$

$$= \frac{1}{\mu_0} \int dA [\mathbf{E}_I \times (-\nabla\chi) - \nabla\psi \times (-\nabla\chi)] \cdot \hat{\mathbf{z}} \tag{12}$$

$$= \frac{1}{\mu_0} \int dA S_z^{\text{pot,I}} + \frac{1}{\mu_0} \int dA (\nabla\psi \times \nabla\chi) \cdot \hat{\mathbf{z}}. \tag{13}$$

Now the latter term that depends on the non-inductive component is zero:

$$\int dA (\nabla\psi \times \nabla\chi) \cdot \hat{\mathbf{z}} = \int dA [\nabla \times (\psi \nabla\chi) - \psi \nabla \times \nabla\chi] \cdot \hat{\mathbf{z}} \tag{14}$$

$$= \int dA [\nabla \times (\psi \nabla\chi)] \cdot \hat{\mathbf{z}} = \oint \psi \nabla\chi \cdot d\mathbf{l} \tag{15}$$

$$= \oint \psi \frac{\partial\chi}{\partial s} ds = \oint \psi B_s^{\text{pot}} ds = 0. \tag{16}$$

The final line integral over the boundary of the photospheric region vanishes if we assume that the fields are sufficiently localized, in which case we can approximate the ψB_s^{pot} to be small at the boundary (Welsch, 2006). Moreover, depending on the method used to compute the potential field, $B_s^{\text{pot}} = 0$ at the boundary may apply by definition (e.g. Seehafer, 1978).

In order to optimize the total (and free) magnetic energy $E_m(t)$, an a priori estimate for it is required. For the NOAA AR 11158 dataset (Section 2.5) used in this paper, we extracted a reference estimate for the magnetic energy injection $E_m(t)$ from the dataset of Kazachenko *et al.* (2015) (hereafter K2015). K2015 used the PDFI electric field inversion method, which, as discussed in the introduction and Section 2.3, is one of the best-performing inversion methods when tested against simulation data in which the true electric field is known. Therefore we here consider it as the state-of-the-art reference, but other estimates might be used as well (see Section 4 for further discussion). We downloaded the dataset of K2015 from <http://cgem.ssl.berkeley.edu/> and calculated $E_m(t)$ using the Poynting flux maps and the mask provided in the dataset. We then created an ensemble of electric

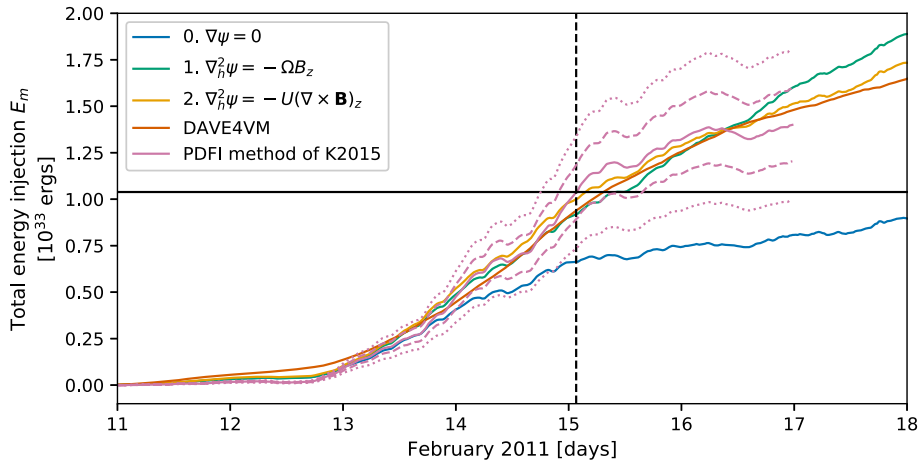


Figure 2 Time evolution of the total magnetic energy $E_m(t)$ injected through the photosphere in AR 11158 based on different electric field estimates. The *black solid line* marks the reference value derived using PDFI electric fields from [K2015](#) on February 15 01:36 UT (denoted by the *vertical dashed line*), which was the closest time before the onset of the strongest flare of the active region at 01:44 UT. The *dashed and dotted purple lines* represent the lower and upper error bars of the [K2015](#) dataset, respectively (see text for details).

field inversion time series using ELECTRICIT (as explained in Section 2.5) and manually searched for optimal values of Ω and U that visually best matched the time evolution of $E_m(t)$ extracted from the [K2015](#) data. After finding the visually best-matching set of values, we optimized U and Ω further by minimizing the root-mean-square error (RMSE) between $E_m(t)$ from [K2015](#) and our data.

3. Results

3.1. Injection of Magnetic Energy

As detailed in Section 2.6, the free parameters of our electric field inversion methods (Ω and U , Section 2.3) were optimized so that the time evolution of the total magnetic energy injection through the photosphere $E_m(t)$ follows the reference result of [K2015](#) as closely as possible. The optimal values searched in this were found to be

$$\begin{cases} \Omega = (21 \pm 0.5)/256 \times 2\pi \text{ day}^{-1} \approx (0.082 \pm 0.002) \times 2\pi \text{ day}^{-1} \\ U = (42 \pm 0.5) \text{ m s}^{-1}, \end{cases} \tag{17}$$

where the error bars correspond to half of the grid spacings of the Ω and U discrete interval used for the final minimization of RMSE between our energy injection estimate and the [K2015](#) result. Even though the time integration in deriving the energy injection (Equation 5) started on February 10 14:00, the RMSE between the energy injections was minimized over the interval February 13 00:00 UT–February 16 23:36 UT. This excludes the first two and a half days of the time series of [K2015](#), over which the energy injection was negligible (see Figure 2).

In addition to these two optimized estimates, we also created an electric field time series using Assumption 0 as well as the DAVE4VM velocity inversion (Sections 2.3 and 2.5). The

evolution of the energy injection for all four estimates and the reference [K2015](#) result are plotted in [Figure 2](#) for the period February 11 – 18, 2011. As seen from the figure, the magnetic energy evolution follows a general trend, in which first a clear increase in the energy injection rate occurs around February 13 (which coincides with the steepest flux emergence rate of the active region, see [Figure 4](#)). The increase continues until the X-class flare (onset on February 15 01:44 UT), after which the [K2015](#) estimate (purple curve) shows a slight decrease in the energy injection rate, which is also discernible in some of our own estimates. The estimate with vanishing non-inductive component (Assumption 0, blue curve) clearly underestimates the energy injection when compared to the other estimates (*e.g.* underestimation of [K2015](#) by 36% at the time of the flare), but reproduces the overall trend of the energy injection rate before and after the flare quite well. Assumptions 1 and 2 (green and orange curves) for the non-inductive component give a roughly similar time evolution for $E_m(t)$ when compared to each other, and they are also consistent with the DAVE4VM (red curve) and [K2015](#) results. There are, however, some obvious differences. The energy injection rate based on Assumption 1 does not decrease as in the other estimates, which causes a clear overestimation of the energy injection from February 16 onward. On the other hand, the energy injection is underestimated between February 14 and 16. The estimate based on Assumption 2 is more consistent with the results of [K2015](#), underestimating its result only slightly near the time of the flare and overestimating it slightly near the end of the [K2015](#) time series (February 16 23:36 UT). There is also a good consistency between Assumption 2 and the DAVE4VM estimate despite the very different electric field inversion methods.

[K2015](#) provided two estimates for the relative error in the vertical Poynting flux S_z and consequently also in the energy injection $E_m(t)$: the first, lower estimate is $\pm 14\%$ and includes only the error arising from the noise in the input vector magnetograms, whereas the second, nominal estimate $\pm 29\%$ also includes the method-related uncertainties of their PDFI electric field inversion method ([Kazachenko, Fisher, and Welsch, 2014](#)). We consider the lower estimate more important because the method-related uncertainties can be assumed to be larger for our electric field estimates, which are partly based on *ad hoc* assumptions. As illustrated in [Figure 2](#) from February 14 00:00 UT until February 16 23:36 UT (the end of the [K2015](#) time series), our energy injection estimates fall well inside the bounds of the upper error estimate and almost within the lower estimate as well. More precisely, the relative difference between any of our estimates and the [K2015](#) estimate is smaller than 15% over this interval. Before February 14, larger relative differences are found, but since the values of the injected energy are small before this time, these correspond to small absolute differences. Hereafter we mostly disregard this interval and discuss results only after February 14 00:00 UT.

Apart from Assumption 0, our energy injection estimates (Assumptions 1 and 2, and DAVE4VM) differ by less than 19% from each other from February 14 until the end of the time series (February 18 00:00 UT).

Our results also reproduce the consistency between the [K2015](#) and DAVE4VM estimates (within both error bounds of [K2015](#)) that has been found by [K2015](#), who compared their $E_m(t)$ estimate to the DAVE4VM estimates of [Liu and Schuck \(2012\)](#) and [Tziotziou, Georgoulis, and Liu \(2013\)](#). Our DAVE4VM estimate is completely independent of these previous studies, with differences both in the input vector magnetograms fed to DAVE4VM and in the way DAVE4VM is executed ([Section 2.3](#)). From the two previous DAVE4VM studies of AR 11158, the study by [Liu and Schuck \(2012\)](#) is more consistent with ours, whereas [Tziotziou, Georgoulis, and Liu \(2013\)](#) used a very different approach in projecting the data to a local Cartesian coordinate system. Our estimate of the energy injection overestimates the result of [Liu and Schuck \(2012\)](#) (see their [Figure 14](#)) by $\sim 5\%$ at reference times February 15 00:00 UT and February 16 $\sim 18:00$ UT (the end of the [Liu and Schuck, 2012](#) time

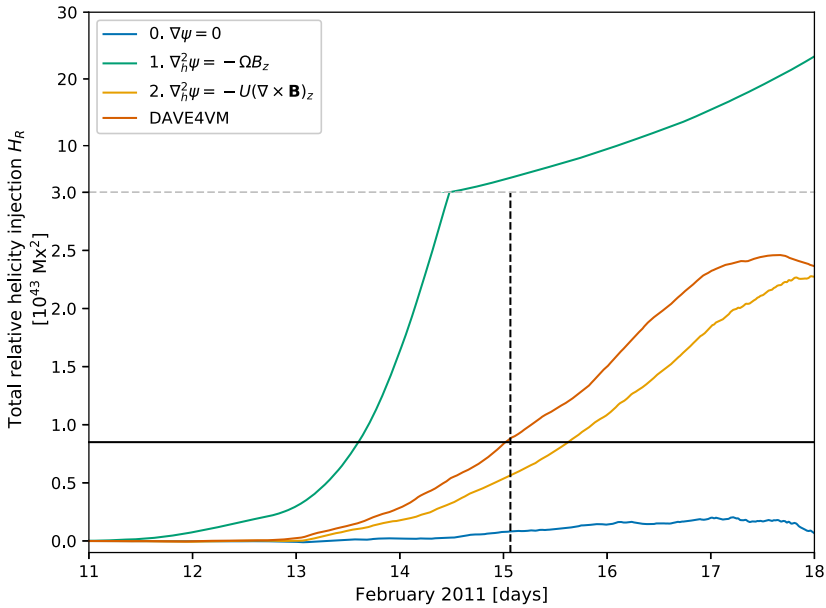


Figure 3 Time evolution of the relative magnetic helicity $H_R(t)$ injected through the photosphere in AR 11158 based on different electric field estimates. The horizontal black solid line marks the reference value derived using PDFI electric fields from K2015 on February 15, 01:36 UT (denoted by the dashed line), near the onset time of the strongest flare of the active region. The horizontal dashed gray line marks the point where the scaling of the y-axis changes to better include the helicity injection based on Assumption 1 to the plot.

series), and the difference is well within the error bounds $\pm 23\%$ reported by the authors. We overestimate the values of Tziotziou, Georgoulis, and Liu (2013) (see their Figure 5) by 20% and 54% at the reference times above. To make a consistent comparison to these previous studies, we started the time integration of the energy injection on February 12, 00:00 UT, when calculating the differences above.

3.2. Injection of Relative Magnetic Helicity

We also studied the injection of relative magnetic helicity through the photosphere $H_R(t)$ (see Section 2.4) using all four of our electric field series and the optimized values of Ω and U for the electric fields based on Assumptions 1 and 2 (Section 2.3). The results are plotted in Figure 3 with one reference data point from K2015. The figure clearly shows that the helicity injection has much larger differences between the different electric field estimates than in the case of energy injection (Figure 2). Assumption 0 again clearly underestimates the result when compared to the K2015 helicity injection estimate (by $\sim 90\%$ at the time of the flare) as well as compared to the other estimates. Despite success in optimizing the magnetic energy injection for Assumptions 1 and 2 (Section 3.1), the relative helicity injection is not reproduced well for either of the assumptions. Assumption 1 heavily overestimates the K2015 helicity injection estimate (by a factor of 6 at the time of the flare and by a factor of 8 on February 16 00:00 UT). Assumption 2 matches the output of K2015 and our DAVE4VM inversion result much better. The overall trends are similar, but there is now a consistent underestimation (by 34% at the time of the flare and by 13% on February 16 at

00:00 UT). DAVE4VM again produces results similar to K2015: our DAVE4VM estimate is only 4% higher than the result of K2015 at the time of the flare, and on February 16 00:00 UT, it is 20% higher. Comparison to the results of previous DAVE4VM studies by Liu and Schuck (2012) and Tziotziou, Georgoulis, and Liu (2013) on February 15 00:00 UT and February 16, 00:00 UT reveals a 45% and a 37% overestimation for the former (see their Figure 12), and an 8% and a 4% overestimation for the latter (see their Figure 5). Larger differences than those obtained in the case of magnetic energy injection (Section 3.1) are not surprising, considering that now, in addition to the differences in the input data and DAVE4VM setup (Section 2.3), the computation methods of the vector potential A_p are also different: these authors employed a fast Fourier transform (FFT) based method to solve A_p directly from the B_z distribution, whereas we employed the PTD of the magnetic field solved using a finite-difference method, similarly to K2015 (see Appendix B for details).

3.3. Effects of Masking and the AR Tracking Method

3.3.1. Effect of Masking the Noise-Dominated Pixels

As discussed in Section 2.2, we masked pixels to zero when $B < 250 \text{ Mx cm}^{-2}$ in the pixel for any of the three frames that participated in the electric field inversion. The threshold was taken from K2015 to ensure maximal compatibility to their results. On the other hand, according to the noise mask of the SDO/HMI vector magnetograms (Hoeksema *et al.*, 2014), the noise level of the total magnetic field strength in vector magnetograms may be as low as 60 Mx cm^{-2} depending on the orbital velocity of the SDO spacecraft and the position of the observation on the solar disk. Moreover, our noise estimate for the magnetic field components (B_x, B_y, B_z) in the AR 11158 time series created in ELECTRICIT (Section 2.5) gives maximum noise levels of $(\sigma_{B_x}, \sigma_{B_y}, \sigma_{B_z}) \approx (65, 85, 65) \text{ Mx cm}^{-2}$ for the time interval February 10 14:00 UT–February 18 00:00 UT. We determined the noise levels for each magnetogram in the series by fitting a Gaussian to the weak-field core ($B < 300 \text{ Mx cm}^{-2}$, Hoeksema *et al.*, 2014) of the magnetogram pixels (Kazachenko *et al.*, 2015; Welsch, Fisher, and Sun, 2013; DeForest *et al.*, 2007). These noise estimates imply that the “pure noise” magnetic field value is $B \sim 125 \text{ Mx cm}^{-2}$. Therefore, when using the threshold of 250 Mx cm^{-2} in masking, we also removed pixels that carry a significant signal, particularly in the relatively low-noise B_z component. Thus, it is natural to ask whether changing the mask threshold affects our results, particularly the magnetic energy injection estimate used in the optimization of the free parameters Ω and U . In order to quantify this, we repeated our analysis for AR 11158 using two different masking thresholds: 1) 150 Mx cm^{-2} , and 2) 300 Mx cm^{-2} . The first threshold corresponds to the nominal threshold used to differentiate between strong and weak field pixels in HMI vector magnetogram pipeline (Liu *et al.*, 2017), and it is slightly above our own noise estimate of 125 Mx cm^{-2} . Thus, we can assume that the threshold masks at least all pure noise pixels. The second case corresponds to the noise threshold used by Liu and Schuck (2012) in their DAVE4VM study of AR 11158 (Y. Liu 2017, private communication).

Changes in the total unsigned magnetic flux at the active region for different masks are illustrated in Figure 4. Threshold 150 Mx cm^{-2} increases the unsigned flux with a median change of 8%, whereas threshold 300 Mx cm^{-2} introduces a median decrease of 4% over the time interval February 10 14:00–February 18 00:00 used in our analyses.

Changes in the injection of the magnetic energy with the alternative masks are illustrated in Figure 5a. We find a consistent trend that the lower the masking threshold, the higher the energy injection. When the value of the masking threshold is varied, the energy injection for the three ELECTRICIT inversions (Assumption 0–2, Section 2.3) increases/decreases

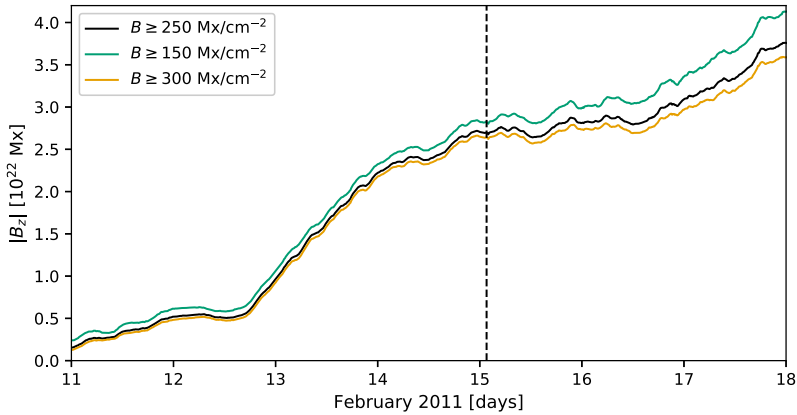


Figure 4 Evolution of the total unsigned $|B_z|$ magnetic flux integrated over the magnetogram surface for NOAA AR 11158 and the three masking thresholds considered here (see text for details).

as much as 21% between February 14 and 17, the maximum median difference being 15%. The estimate based on Assumption 0 is affected more than those based on Assumptions 1 and 2, for which both the maximum and median relative differences to the default case are $\sim 20\%$ and $\sim 12\%$. The strongest change (increase) is found when the mask threshold of 150 Mx cm^{-2} is used. Varying the mask changes the optimized estimates of the energy injection (Assumptions 1 and 2) just within the maximal error bounds of $E_m(t)$ of K2015 ($\pm 29\%$), but not within the lower error bounds ($\pm 14\%$).

The estimate based on the DAVE4VM inversion is practically independent of the choice of the masking threshold. When the mask is varied or even removed entirely, the maximum relative difference between the energy estimates is 4% between February 14 and 17 (the median relative difference is smaller than 2%). Therefore, the DAVE4VM estimate is not included in Figure 5a. It is important to note that the magnetograms are not masked when fed into DAVE4VM, but the mask is applied only afterward. Thus, the fact that DAVE4VM gives practically the same result with and without masking implies that the velocity inversion in DAVE4VM cannot retrieve any significant vertical Poynting flux signal (negative or positive) from the pixels with a low signal-to-noise ratio. The same does not apply to ELECTRICIT inversions: removing the mask entirely increases the (time-integrated) energy injection even as much as 80%. A part of this energy flux signal arises from the use of the potential acute angle disambiguation in the weak-field region (Section 2.2), since using the alternative methods (radial acute angle or random disambiguation) results in a smaller increase ($< 40\%$) when the mask is removed entirely. Still, as we argued in Section 2.2, the choice of the weak-field disambiguation method does not affect the results when any of the three masking thresholds considered above are applied. The reason for the spurious energy injection signal arising from the use of the potential acute angle disambiguation is likely the large-scale “shadow-like patterns” (Liu *et al.*, 2017) introduced to the azimuth by the method, which somehow interplay with the large-scale electric field to produce a significant vertical Poynting flux signal. Nevertheless, removing the mask entirely introduces a clear increase in the energy injection regardless of the weak field disambiguation method, which implies that there are also other method-related factors in ELECTRICIT inversions that enable retrieval of a significant vertical Poynting flux signal from the noise-dominated pixels. The fact that this is not seen in DAVE4VM output implies that the signal is at least partly spurious, thus emphasizing the need for using the mask.

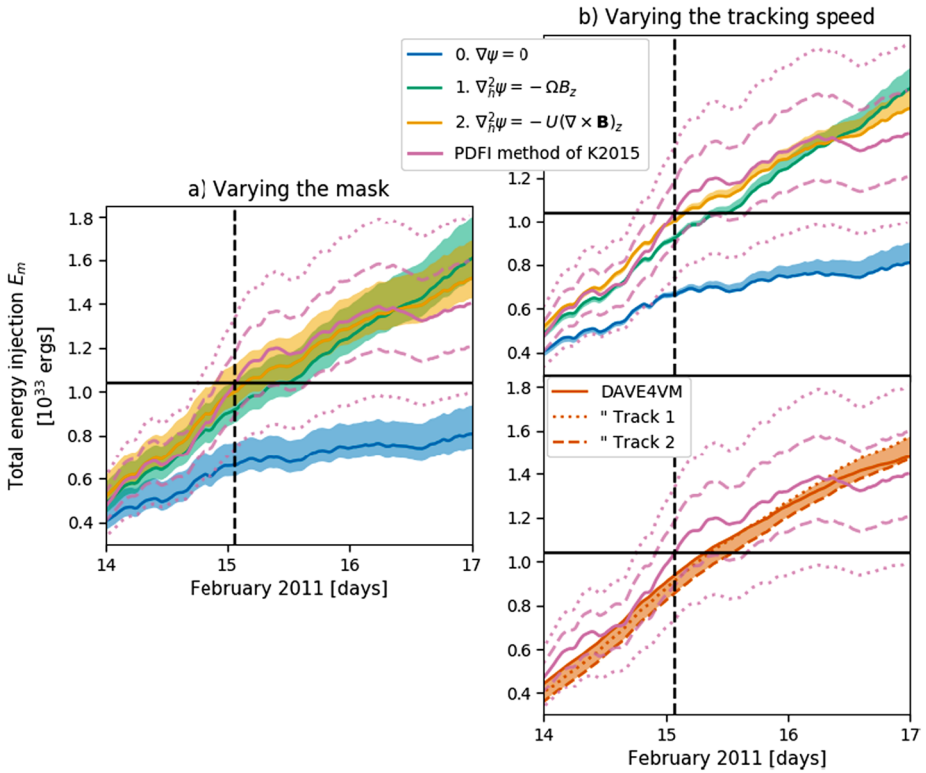


Figure 5 Total magnetic energy $E_m(t)$ injected through the photosphere in AR 11158 and the variations arising from changes in the **a)** masking threshold and **b)** tracking speed of the active region (see text for details). The shaded region around each curve covers the variation in the energy injection due to the respective changes in the data processing. Dashed purple lines give the upper and lower error bars of the K2015 result. Track 1 refers to the tracking made using the same differential rotation profile as in the SHARP data, and Track 2 refers to our default tracking reinforced with coalignment. The black dotted line marks the reference value from K2015 on February 15 01:36 UT (denoted by the dashed line) near the onset time of the strongest flare of the active region.

For the relative helicity injection, we find similar trends when varying the masking thresholds as for the energy injection: *i.e.* lowering the threshold increases the injected relative helicity. Relative changes between the estimates using the alternative masking thresholds are strongest for the estimate based on Assumption 0 (maximum and median differences to the default are 118% and 41% between February 14 and 17), but the absolute changes are small and the large relative changes are simply due to the small absolute values of $H_R(t)$. For Assumptions 1 and 2, the maximum and median relative changes are 10% and 8%, and 19% and 10%. For the DAVE4VM estimate, the maximum and median relative changes are 13% and 8%, respectively. These changes do not alter the basic findings made in Section 3.2 where the default mask was used.

3.3.2. Effect of the Tracking Speed of the AR

Another data processing parameter that likely affects the results is the speed at which the active region is tracked during its solar disk passage. Our choice for the tracking speed is

based on the “2-day lag” differential rotation profile of Snodgrass (1983), but other options are also possible. In the SDO/HMI SHARP data product, the active regions are tracked using a different profile (Hoeksema *et al.*, 2014), which gives a $0.27 \text{ deg day}^{-1}$ ($\sim 2 \text{ km s}^{-1}$) faster speed for the center of AR 11158 along the line of constant latitude. On the other hand, Kazachenko *et al.* (2015) and Welsch, Fisher, and Sun (2013) tracked the active region using a coalignment procedure that minimizes the whole-frame shifts between successive frames of the magnetogram time series (in the local Cartesian coordinate system). In practice, this approach results in a temporally variable tracking speed and also includes motions along the meridional direction. Owing to the velocity in the ideal Ohm’s law, it is expected that changing the tracking speed, and thus also the velocity of the rest frame, has an effect on the electric field and thus also on the energy and relative helicity injections.

In order to quantify how sensitive our results are to the choice of the tracking speed, we repeated our analysis for AR 11158 using two alternative tracking schemes: 1) the differential rotation profile employed for the SHARP data product, and 2) our default differential rotation profile reinforced by coalignment, which minimizes the whole-frame shifts between the successive frames. For the coalignment we used the method presented by Welsch, Fisher, and Sun (2013) in which the coalignment is achieved using low-noise LOS magnetograms. For each time step, we interpolated an SDO/HMI full-disk LOS magnetogram to the local Cartesian grid determined by our default tracking procedure (Section 2.2 and Appendix A.1). Then the whole-frame shift to the previous interpolated LOS data frame was evaluated, and the local Cartesian grid was shifted so that it canceled the measured whole-frame shift. This shifted grid was then used as the final grid to which the vector (and LOS) magnetograms were interpolated at that time. For our 11158 time series, the coalignment results in a tracking speed that is on average $0.28 \text{ deg day}^{-1}$ faster than our default case (with standard deviation of $0.21 \text{ deg day}^{-1}$).

As illustrated in Figure 5b, the effect of changing the tracking speed has only a small effect on the estimate of the energy injection when the three ELECTRICIT inversions (Assumptions 0–2) are considered, and the changes are most of the time within the lower error bars ($\pm 14\%$) of the K2015 estimate. The first tracking scheme (SHARP rotation profile) produced a small increase toward the end of the time series for all three estimates, the maximum and the median relative changes being 11% and 4% for Assumption 0, and 6% and 3% for Assumptions 1 and 2 between February 14 and 17. The second tracking scheme (default rotation profile and coalignment) caused only a marginal change: the maximum and the median relative differences to the default cases were at most 6% and 2% over all three assumptions. When the DAVE4VM estimate was considered, the tracking speed had a more significant effect. When we used the SHARP rotation profile, the default energy injection was first underestimated, but then overestimated toward the end of the time series (the maximum and median relative differences are 10% and 3%), whereas for the coaligned case, there is a consistent decrease in the injection throughout the time series (the maximum and median relative differences are 19% and 7%). Despite the relatively large effect on the energy injection, the DAVE4VM estimates are still most of the time within the lower error bars of K2015. The coalignment brings the DAVE4VM estimate closer to the estimate by Liu and Schuck (2012) (see Section 3.1): now the estimates differ by only 1% near the time of the flare on February 15 00:00 UT (but still by 5% on February 16 18:00 UT).

The effect of the tracking speed on the estimate of relative helicity injection does not affect the main findings made in Section 3.2 where the default tracking speed was used. When the tracking speed is varied, the estimate based on Assumption 0 exhibits large relative changes ($> 100\%$) because of its low absolute values, but the other estimates change by 8% at most.

4. Summary and Discussion

We examined the magnetic energy and relative helicity injection through the photosphere in NOAA AR 11158 using an ensemble of estimates for the photospheric electric field. We employed our own data processing and electric field inversion toolkit, ELECTRICIT, to create a vector magnetogram time series centered at the active region. We further created four time series of photospheric electric field maps from this data. One series was created using the DAVE4VM velocity inversion code, and the three other series were created within ELECTRICIT. In the latter approach we decomposed the electric field into inductive and non-inductive components. We constrained the inductive component to be exactly consistent with Faraday's law and the time evolution of B_z in our vector magnetogram time series. Since constraining the non-inductive component properly from the measurements requires advanced techniques that are not yet implemented in our toolkit, we used three alternative *ad hoc* assumptions instead. This approach also enabled a straightforward inversion procedure without the need of acquiring and calibrating any photospheric plasma velocity estimates. Our zeroth assumption simply neglects the non-inductive component (setting it to zero), whereas the other two constrain it by setting the horizontal divergence of the electric field proportional either to the vertical magnetic field B_z or the current density j_z . The proportionality constants in the latter two assumptions are free parameters, and we optimized them (*i.e.* chose temporally and spatially constant values for them) so that the resulting electric field time series reproduced the time evolution of the total injection of magnetic energy from the photosphere to the corona $E_m(t)$ as closely as possible (computed as the area- and time-integrated vertical Poynting flux). Using analytical arguments, we showed that this approach also implicitly optimizes the injection of the free magnetic energy. The reference for the optimal $E_m(t)$ was extracted from the results of Kazachenko *et al.* (2015) (here abbreviated as K2015), who used state-of-the-art methods to constrain the non-inductive component.

Using the optimization scheme for the free parameters of the two *semi-ad hoc* electric fields above, we are able to reproduce the time evolution of the injected magnetic energy $E_m(t)$ well within the upper error bounds of the result of Kazachenko *et al.* (2015), and almost (exceeding only by 1%) within the lower error bounds, for the time period when photospheric energy injection had reached significant values (more than half of the values at the time of the strongest X-class flare of the active region). We consider the lower error bounds more important here, since they only include the error arising from the noise in the vector magnetograms, whereas the upper estimate also includes method-related uncertainties, which can be safely assumed to be much higher for our (*semi-ad hoc*) estimates. Our DAVE4VM study of the energy injection is consistent with the results of Kazachenko *et al.* (2015) and with the previous DAVE4VM-based results for NOAA AR 11158 by Liu and Schuck (2012) within their error bounds. However, the fourth electric field time series of our ensemble, derived setting the non-inductive component to zero (our zeroth *ad hoc* assumption) clearly underestimates both the total (and thus also the free) magnetic energy injection. This is in line with previous simulation studies by Fisher, Welsch, and Abbett (2012) and Kazachenko, Fisher, and Welsch (2014), who showed that by assuming a vanishing non-inductive component, one clearly underestimates the vertical Poynting flux. We also find that the assumption underestimates the relative helicity injection, which is again consistent with the previous results (*e.g.* Kazachenko, Fisher, and Welsch, 2014). However, the positive sign of the total relative helicity injection is reproduced correctly.

The underestimation of both magnetic energy and relative helicity injection implies that TMF simulation studies in which the photospheric boundary condition directly or indirectly employs Assumption 0 likely underestimate the budgets of these quantities in the simulation domain (although one should keep in mind that the coronal budget of these quantities

cannot be deduced solely from the photospheric fluxes, as discussed further below). Despite this implication, Gibb *et al.* (2014) successfully simulated the formation of a flux rope in NOAA AR 10977, yielding results that were partly consistent with the observed X-ray sigmoid, even though the photospheric boundary condition was based on Assumption 0. However, their simulation did not include the eruption of the flux rope. Owing to the central role of (free) magnetic energy and possibly also relative helicity in flux rope ejections, underestimating these quantities may be important only when the lift-off and subsequent eruption dynamics are considered. This is implied by the results of Cheung and DeRosa (2012), whose TMF simulation of NOAA AR 11158 gave only a “quiescent” coronal evolution when Assumption 0 was used, whereas the use of Assumption 1 resulted in a series of flux rope ejections.

The fact that we are able to reproduce the time evolution of the magnetic energy injection well using Assumptions 1 and 2 implies that both of these *semi-ad hoc* electric field inversion methods are flexible enough for estimating the total (and free) magnetic energy injection for AR 11158 considered here. Although AR 11158 contains many of the basic properties of a typical solar active region and we managed to analyze its evolution from its initial emergence well beyond its strongest eruption, we still cannot make far-fetched conclusions about the applicability of our approach to other active regions. Moreover, it should be noted that our approach does not constrain the free parameters Ω and U universally, but only presents an approach for finding the best values for them by optimizing the photospheric energy injection over some prescribed time interval. Thus, the optimal values for these parameters do not depend only on the properties of the active region in question, but also on the choice of the optimization interval. For AR 11158, we focused on the interval around the eruption time of the strongest flare of the region (our optimization interval was February 13 00:00 UT – 16 23:36), but if earlier or later eruptive activity of the region were to be studied, another choice of the interval might be required. Further quantification of these questions would require a statistical study where our optimization approach would be applied to several active regions with variable properties.

Another limitation in our approach is its requirement for a reference energy injection estimate for optimization of the free parameters Ω and U . For an arbitrary active region, reference estimates (such as the PDFI estimate by K2015, used here) are not readily available. One straightforward way to acquire such an estimate is to use DAVE4VM: for AR 11158, it would give almost the same optimization result as the PDFI method in K2015. Although the electric fields derived using DAVE4VM are suitable for deriving such a reference estimate, they are, however, unsuited for driving of coronal simulations, since the electric fields do not fulfill Faraday’s law (see Section 2.3 for details). If one cannot acquire an estimate for the whole time evolution of the energy injection, the reference for the optimization procedure could include an energy estimate only at some given time(s). For instance, for AR 11158, an estimate of the coronal energy content just before the strongest (X-class) flare of the active region could be used. Such an instantaneous estimate could be acquired *e.g.* from a non-linear force-free field (NLFFF) extrapolation. Although we know that the energy injection through the photosphere is not equivalent to the energy stored in the simulated coronal domain, Kazachenko *et al.* (2015) found that the (ordinary) NLFFF extrapolation estimates of the coronal energy content were consistent with the photospheric injection within error bounds in AR 11158.

Our optimal values of the free parameters Ω and U are lower than the values used by the original authors who proposed these *semi-ad hoc* estimates (Section 2.3). Cheung and DeRosa (2012) tested two non-zero values for Ω , $0.125 \times 2\pi$ and $0.250 \times 2\pi \text{ day}^{-1}$, from which the latter was primarily used. The values are 50% and 200% higher, respectively than

our value $\Omega = 0.082 \times 2\pi \text{ day}^{-1}$. Cheung *et al.* (2015) tested two values for U , 1100 and 2200 m s^{-1} , from which the former was primarily used in the paper. Both values are almost two orders of magnitude higher than our value $U = 42 \text{ m s}^{-1}$. When we ran our analysis using the nominal values of the previous authors $\Omega = 0.25 \times 2\pi \text{ day}^{-1}$ and $U = 1100 \text{ m s}^{-1}$, the magnetic energy and relative helicity injection were overestimated by factors 1.4 and 18 for the former case and by 9 and 15 for the latter at the time of the strongest X-class flare (when compared to the best reference values considered in this paper). Thus, our results imply that the TMF simulation of AR 11158 by Cheung and DeRosa (2012) may have had slightly overestimated total and free magnetic energy budgets and a heavily overestimated relative helicity budget. As opposed to the TMF simulation of Cheung and DeRosa (2012) that included the entire AR 11158, the simulation of Cheung *et al.* (2015) focused on homologous helical jets emanating from a small subregion of NOAA AR 11793. Thus, further conclusions about the realism of their magnetic energy and relative helicity budgets would require an additional study of this particular subregion of AR 11793. Moreover, we reiterate that we must be careful when drawing conclusions from the photospheric injections on the realism of the magnetic energy and relative helicity budgets in the simulated coronal domain. This is because part of the photospheric injection may be removed through the other boundaries of the simulation domain or may be transformed into other forms in processes within the domain. Moreover, for the special case of the magnetic energy injection in TMF simulations, not even the photospheric injections are equivalent between our estimate and the corresponding TMF simulation: we derived the photospheric Poynting flux using the horizontal components of the magnetic field from a vector magnetogram, whereas in a TMF simulation, these components are determined by the magnetofrictional evolution in the coronal domain.

A more comprehensive analysis of the relationship between the photospheric fluxes and coronal budgets would require additional (TMF) simulation studies, which is beyond the scope of this paper. For NOAA AR 11158, we can make some further conclusions by comparing our fluxes to the simulation energy budgets reported by Cheung and DeRosa (2012). Their reprojected magnetograms and the resulting electric field inversion were very similar to that of this study, except for the fact that they used LOS magnetograms. Using the functionality in ELECTRICIT, we created an additional B_z magnetogram time series from full-disk LOS magnetograms to the same grid as our nominal vector magnetogram time series (Section 2.5), inverted the horizontal electric field E_h using Assumption 1 and $\Omega = 1/4 \times 2\pi \text{ day}^{-1}$, and estimated the energy injection using B_h from our nominal vector magnetogram series. Near the time of the X-class flare, our photospheric injection estimate was only 13% higher than the total energy content reported by Cheung and DeRosa (2012), so at least in terms of energy, the photospheric injection seems to match the coronal budget. Using this LOS magnetogram dataset, we can also estimate the relative helicity injection, which is still overestimated by a factor of ~ 16 at the time of the strongest flare of the active region. Even though the dissipation of relative helicity in the corona can be considered negligible (*e.g.* Pariat *et al.*, 2015), part of the helicity injected from the photosphere can and will be removed through the boundaries of the domain, particularly during CME flux rope ejections (*e.g.* Pariat *et al.*, 2017), in which case the flux rope also carries part of the helicity away from the domain. Jing *et al.* (2012) found in their NLFFF extrapolation study of AR 11158 that the total coronal relative helicity content was substantially lowered in tandem with the strongest M6.6 and X2.2 class flares (and the related flux rope ejections) in the active region. However, the magnitude of helicity removals (altogether $\sim 5 \times 10^{42} \text{ Mx}^2$) is not sufficient to remove the overestimated photospheric relative helicity injection in the Cheung and DeRosa (2012) simulation implied by our results. This finding poses the interesting question whether the strong activity in the simulation (several flux rope ejections over

two days) is partly driven by this heavy overestimation of the relative helicity budget in the corona.

Despite the relatively good optimization of the magnetic energy injection, the optimal values of Ω and U only poorly reproduce the injection of relative magnetic helicity through the photosphere. Even with an energy-optimized value of Ω , Assumption 1 results in a heavy overestimation of the relative helicity injection. This is no surprise when we take a closer look at the calculation of the electric field and the vector potential A_p . First, we write the total flux of relative helicity arising from the non-inductive component dH_R^{NI}/dt for a given time using the poloidal potential P and the non-inductive potential ψ :

$$\begin{aligned} \frac{dH_R^{NI}}{dt} &= -2 \int dA [A_p \times (-\nabla\psi)] \cdot \hat{\mathbf{z}} = -2 \int dA [(\nabla \times P\hat{\mathbf{z}}) \times (-\nabla\psi)] \cdot \hat{\mathbf{z}} \\ &= 2 \int dA [\partial_x P \partial_x \psi + \partial_y P \partial_y \psi]. \end{aligned} \tag{18}$$

From the definition of Assumption 1 (Section 2.3) and the PTD of the magnetic field (Appendix B), we obtain

$$\nabla_h^2 P = -B_z = \Omega^{-1} \nabla_h^2 \psi. \tag{19}$$

Since we employ the same grid, numerical solver, and homogeneous Neumann boundary conditions for solving both the P and ψ potentials, they are the same potential up to a constant scale factor $\psi = \Omega P$. Thus we obtain

$$\frac{dH_R^{NI}}{dt} = 2\Omega \int dA [(\partial_x P)^2 + (\partial_y P)^2] \geq 0. \tag{20}$$

We can immediately see that for any non-zero Ω we always obtain a positive contribution to the relative helicity injection from the non-inductive component of the electric field. On the other hand, when we use Assumption 2 to estimate the non-inductive component, the ψ and P potentials decouple and thus no immediate positive bias is introduced to the relative helicity injection. Assumption 2 with the energy-optimized value of U indeed reproduces the time evolution of the relative helicity injection much better, having a similar temporal trend as the K2015 and our DAVE4VM estimates (but exhibiting a consistent underestimation, see Section 3.2).

It is not a surprise that neither of the *ad hoc* Assumptions 1 or 2 are able to reproduce the injection of both the magnetic energy and relative helicity accurately. Still, the results clearly show that Assumption 2 succeeds much better than Assumption 1 in this sense. This implies that Assumption 2 is more suitable for practical use, for example when the electric fields are used as a boundary condition in coronal modeling. On the other hand, Assumption 1 offers an interesting way to study the effect of overestimating the photospheric helicity injection while still producing approximately correct energy injection in data-driven (TMF) simulations. For example, a comparative TMF simulation study using electric field boundary conditions based on (optimized) Assumptions 1 and 2 would be revealing in this sense.

Finally, we also studied the effect of some of the data processing methods on the results, in particular on the magnetic energy injection. These include the masking threshold used to remove noise-dominated pixels from the data and the tracking speed of the active region, both of which have so far been poorly covered in the literature. We tested three masking thresholds that correspond to approximately 1.5, 2.5, and 3 times the nominal noise level of the total magnetic field B in SDO/HMI vector magnetograms (100 Mx cm^{-2}), the second of which was the default used in this paper. We find that the lower the threshold, the

higher the magnetic energy and relative helicity injections. The changes introduced to the energy injection by varying the mask are within the upper error bars of our reference energy injection estimates, but clearly above the lower error estimate that we mainly considered here. Thus, in order to reach consistency with the **K2015** reference estimate (or with some other similar reference estimate of the energy injection) approximately within the $\pm 14\%$ error bounds (as we almost managed to do, exceeding this only by 1%), the mask threshold should be chosen consistently with the reference. However, the choice of the mask does not affect the main findings made above. The DAVE4VM estimate for the energy injection is completely insensitive to the choice of mask, even when the mask is removed entirely, implying that DAVE4VM cannot retrieve any significant vertical Poynting flux signal from the noise-dominated weak-field pixels. Energy estimates based on ELECTRICIT inversions in turn increase significantly when the mask is removed. The fact that this increase is only seen in the ELECTRICIT output but not in the DAVE4VM inversions implies that this increase is a spurious result arising from the method-related properties of the ELECTRICIT inversion, thus emphasizing the need for using a mask.

For the tracking speed of the active region, we tested three alternatives and found that our three ELECTRICIT estimates are insensitive to the choice of the tracking speed. However, the tracking speed had a more considerable effect on the DAVE4VM estimate, particularly when the tracking of the active region in vector magnetograms was reinforced using additional coalignment that minimized the whole-frame shifts between successive frames. However, the change was still within the previously reported error bars of the DAVE4VM inversion, and the resulting DAVE4VM energy injection estimate was still most of the time within the lower error bars of the **K2015** estimate discussed above. The stronger effect on DAVE4VM likely arises from the fact that the method explicitly tracks the image motions in vector magnetograms, which is more sensitive to the tracking speed and to the related changes in the velocity of the rest frame. Overall, despite the considerable effects that we found when varying the mask threshold and the tracking speed, none of the effects were sufficiently strong to conclusively allow preferring some of the options over others.

5. Conclusions

By studying the area- and time-integrated Poynting flux, we were able to show that the data-processing and electric field inversion approach presented in this paper, which employs only photospheric magnetic field estimates, is capable of reproducing the time evolution of the total (and implicitly also the free) magnetic energy injected through the photosphere, when compared to recently published estimates for solar active region NOAA 11158. The lack of the photospheric velocity estimates in the approach was compensated for using *ad hoc* assumptions whose free parameters allowed the optimization of the magnetic energy injection. Despite the good reproduction of the magnetic energy injection, the *ad hoc* assumptions reproduced the injection of relative helicity poorly, reaching at best a modest underestimation of the reference values from previous studies. The *ad hoc* assumption that constrains the non-inductive potential by setting the horizontal divergence of the electric field proportional to the vertical current density performed best in this sense, implying that it is the most suitable assumption for practical applications. Finally, we showed that our results are not very sensitive to the choice of the masking threshold for removing the noise-dominated pixels and the tracking speed of the active region. However, our inversion requires some mask for the weak-field pixels, and this mask should be chosen consistently with the reference when optimizing the free parameters of the inversion using the magnetic energy injection. The main datasets used in this paper can be found at <https://zenodo.org/record/1034404>.

Acknowledgements We thank the HMI team for providing us with the LOS and vector magnetic field SDO/HMI data. EL acknowledges the doctoral program in particle physics and universe sciences (PAPU) of the University of Helsinki and the Magnus Ehrnrooth Foundation for financial support. EK and JP acknowledge the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Programme Project SolMAG 4100103. This research has made use of SunPy, an open-source and free community-developed solar data analysis package written in Python (Mumford *et al.*, 2015).

Disclosure of Potential Conflicts of Interest The authors declare that they have no conflicts of interest.

Appendix A: Modifications to the Existing Data Processing Methods

A.1 Tracking Scheme of Active Regions in SHARP Data

As explained by Bobra *et al.* (2014) and Hoeksema *et al.* (2014), reprojected vector magnetograms in SHARP time series are created by tracking each HMI Active Region Patch (HARP) using a fixed rotation rate for the center point of the patch. The rotation rate is determined from the latitude of the center point using a differential rotation profile. Instead of the differential rotation profile described in Hoeksema *et al.* (2014), we employ the “2-day lag” profile from Snodgrass (1983). For the NOAA 11158 vector magnetogram time series used in this paper, our tracking speed is $\sim 2 \text{ km s}^{-1}$ (0.27 deg per day) slower than the SHARP tracking speed. We observe that when using this method, the center of the active region remains slightly better fixed to the center of the patch. As explained by Sun (2013), the final reprojected SHARP magnetogram is created for each frame so that the disambiguated full-disk vector magnetogram is reprojected on the solar surface, *i.e.* interpolated from image pixels to a new grid that corresponds to a Lambert CEA projection of the solar surface covering the predefined active region size. The CEA projection is made so that the active region is “viewed directly from above”, *i.e.* using a heliographic coordinate system where the center of the active region is at the intersection of the central meridian and the equator. The grid spacing is 0.03° in projected heliographic coordinates, corresponding to the SDO/HMI resolution at the disk center. We have modified this procedure so that instead of a CEA projection, we employ Mercator projection, since conformal mappings are preferred when optical flow methods (such as DAVE4VM) are used (Welsch *et al.*, 2009; Kazachenko *et al.*, 2015). After reprojecting the full-disk magnetogram data, we rotate the interpolated magnetic field vectors from the SDO/HMI image basis (B_ξ, B_η, B_z) (uniquely defined by the orthogonal image axes and LOS direction) to the heliographic basis $(B'_r, B'_\theta, B'_\phi)$. The heliographic basis is chosen consistently with the Mercator map projection so that the center point of the active region is at the intersection of the central meridian and the equator. Transformation from the default heliographic basis (B_r, B_θ, B_ϕ) where the true solar equator has latitude $\lambda = 0$ to the patch-centered basis $(B'_r, B'_\theta, B'_\phi)$ can be made using the following transformation (which is not included in the references listed above):

$$\begin{pmatrix} B'_r \\ B'_\theta \\ B'_\phi \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} B_r \\ B_\theta \\ B_\phi \end{pmatrix} \tag{21}$$

$$\alpha(\lambda, \phi, \lambda_c, \phi_c) = \text{sgn}(\lambda_c) \text{sgn}(\phi - \phi_c) \times \cos^{-1} \left[(\cos \lambda_c \cos \lambda + \sin \lambda_c \sin \lambda \cos(\phi - \phi_c)) / \cos \lambda' \right] \tag{22}$$

$$\lambda' = \sin^{-1} \left[\sin \lambda \cos \lambda_c - \sin \lambda_c \cos \lambda \cos(\phi - \phi_c) \right], \tag{23}$$

where λ' is the latitude in heliographic coordinates where the patch center $(\lambda'_c, \phi'_c) = (0, 0)$.

This is different from SHARP data in which the heliographic basis corresponds to Carrington coordinates in which the equator is the true solar equator (Sun, 2013). Finally, we further transform into local Cartesian basis where we treat the projected active region patch as a flat surface:

$$\begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix} = \begin{pmatrix} B'_\phi \\ -B'_\theta \\ B'_r \end{pmatrix} \quad (24)$$

where we drop the apostrophes to simplify notation.

A.2 Algorithm for the Removal of Spurious Flips in the Azimuth of the Magnetic Field

Welsch, Fisher, and Sun (2013) defined spurious flips of the azimuth of the magnetic field ϕ as sudden jumps where ϕ jumps approximately from one disambiguation to the other $\phi \rightarrow \phi + 180^\circ$ and then quickly back. In a time series, this flipping can be seen as rapid blinking in the transverse magnetic field components. The authors proposed a temporal smoothing algorithm to recognize and remove such spurious flips. The algorithm detects these flips using two conditions. First, the change in azimuth between successive frames must be significant enough (over 120°):

$$\|\phi(x_m, y_n, t_k) - \phi(x_m, y_n, t_{k-1})\| > 120^\circ, \quad (25)$$

where $\|x\|$ refers to the absolute acute angle of x :

$$\|x\| = \begin{cases} |x|, & |x| \leq 180^\circ \\ 360^\circ - |x|, & |x| > 180^\circ. \end{cases} \quad (26)$$

Second, if a spurious azimuth flip candidate at t_k is removed by flipping the azimuth back to the “correct” value ($\phi \rightarrow \phi + 180^\circ$), the procedure must decrease the sum of unsigned azimuth differences over the nearby frames:

$$S = \sum_{l=-R}^R \|\phi(x_m, y_n, t_{k+l}) - \phi(x_m, y_n, t_k)\|, \quad (27)$$

i.e. the procedure must smooth the azimuth in the time domain. Here R is the expected length of that spurious flip in time, and it gives the upper limit for the number of frames the spurious flip may last in order to be recognized by the algorithm. R is the only free parameter in the algorithm, and Welsch, Fisher, and Sun (2013) suggested two values for it: $R = 2$ or $R = 4$, of which they used the former. We chose to use $R = 4$ instead, since it produces a more stable evolution of the azimuth in visual inspection. Another major qualitative departure from the work of Welsch, Fisher, and Sun (2013) is that we applied the flip-removal procedure after interpolating the SDO/HMI data to a local Cartesian system in Mercator projection (Appendix A.1), whereas they applied the procedure in native HMI pixels. Thus, in our approach the erroneous values of pixels where the azimuth had a spurious flip propagated in the interpolation. We therefore employed an additional step after recognizing and fixing the spurious flips: we smoothed each fixed pixel and all of its eight neighboring pixels using a Gaussian smoother with $\sigma = 1$ pixel truncated at 1σ as a way to mitigate the effect of propagating the erroneous azimuth to the neighboring pixels in the interpolation.

An additional motivation for this smoothing is that it removes some of the artifacts of the flip-removal procedure itself: the procedure leaves some ragged structures and singular pixels of inconsistent azimuth values to the data, which we consider spurious, particularly when taking into account that the minimum energy disambiguation tries to reduce currents, *i.e.* to smooth the azimuth spatially.

Appendix B: Numerical Implementation of the Electric Field Inversion Methods in ELECTRICIT

In the PDFI method, Faraday’s law (Equation 4) is uncurled using the poloidal–toroidal decomposition of the magnetic field (Fisher *et al.*, 2010):

$$\mathbf{A} = \nabla \times P\hat{\mathbf{z}} + T\hat{\mathbf{z}} \tag{28}$$

$$\mathbf{E}_I = -\nabla \times \dot{P}\hat{\mathbf{z}} - \dot{T}\hat{\mathbf{z}}, \tag{29}$$

where \mathbf{A} is the magnetic vector potential $\nabla \times \mathbf{A} = \mathbf{B}$, and \dot{P} and \dot{T} are the partial time derivatives of two-dimensional poloidal and toroidal potentials, respectively. They can be solved from the Poisson equations of which source terms are determined by the $\partial \mathbf{B} / \partial t$ field. The time derivative of the poloidal potential can be solved from

$$\nabla_h^2 \dot{P} = -\frac{\partial B_z}{\partial t}, \tag{30}$$

where ∇_h^2 is the horizontal Laplacian $\partial_x^2 + \partial_y^2$. We here employed only the horizontal components of the inductive electric field $\mathbf{E}_h^I = (E_x^I, E_y^I)$, which are determined solely from \dot{P} . We solved Equation 30 using the same homogeneous Neumann boundary conditions and the numerical solver (FISHPACK, Swarztrauber and Sweet, 1975) as Kazachenko, Fisher, and Welsch (2014). As noted by Kazachenko, Fisher, and Welsch (2014), FISHPACK modifies the source term of a Poisson equation by adding a constant value to it, if it is found to be inconsistent with the homogeneous Neumann boundary conditions. As a result of this correction, the output electric field is slightly inconsistent with Faraday’s law. To remove this inconsistency, we added a *post facto* correction to the horizontal electric field components, as done by Fisher *et al.* (2010). To further ensure the consistency with Faraday’s law, we also decided to use different finite-difference approximations for Equations 28 and 29 than Kazachenko, Fisher, and Welsch (2014). The FISHPACK solver employs the five-point stencil for the Laplacian, which implicitly assumes that the first-order spatial derivatives of the solution of the Poisson equation are determined in half-grid points. Thus, if the first-order derivatives in Equation 29 and Faraday’s law (Equation 4) for the horizontal components of \mathbf{E}_I are calculated using a typical central difference scheme, $\partial B_z / \partial t$ is not reproduced exactly. Kazachenko, Fisher, and Welsch (2014) argued that the error introduced by this is small and employed a central difference scheme also for the first-order derivatives. However, we wish to reproduce $\partial B_z / \partial t$ from Faraday’s law exactly, and therefore employed the following consistent finite-difference formulas for the spatial derivatives (also used in a similar fashion by Yeates, 2017):

$$E_x^I(x_m, y_{n+1/2}) = -\frac{\partial \dot{P}}{\partial y}(x_m, y_{n+1/2}) = -\frac{\dot{P}(x_m, y_{n+1}) - \dot{P}(x_m, y_n)}{\Delta y} \tag{31}$$

$$E_y^1(x_{m+1/2}, y_n) = \frac{\partial \dot{P}}{\partial x}(x_{m+1/2}, y_n) = \frac{\dot{P}(x_{m+1}, y_n) - \dot{P}(x_m, y_n)}{\Delta x} \tag{32}$$

$$\begin{aligned} \frac{\partial B_z(x_m, y_n)}{\partial t} &= -(\nabla \times \mathbf{E}_I)_z(x_m, y_n) = \frac{\partial E_x^1}{\partial y}(x_m, y_n) - \frac{\partial E_y^1}{\partial x}(x_m, y_n) \\ &= \frac{E_x^1(x_m, y_{n+1/2}) - E_x^1(x_m, y_{n-1/2})}{\Delta y} \\ &\quad - \frac{E_y^1(x_{m+1/2}, y_n) - E_y^1(x_{m-1/2}, y_n)}{\Delta x}. \end{aligned} \tag{33}$$

As can be implicitly read from the equations above, the spatial derivatives consistent with the five-point stencil give the electric field components in a staggered grid (Yee mesh; Yee, 1966) with respect to B_z and $\partial B_z/\partial t$ at (x_m, y_n) . Since many of the TMF simulations employ such a staggered grid (e.g. van Ballegooijen, Priest, and Mackay, 2000; Cheung and DeRosa, 2012), this makes the electric field inversions of ELECTRICIT instantly suitable to be used as the boundary condition data of such simulations, while being simultaneously consistent with the observed time evolution of B_z .

The numerical approach described above was also used to solve the P potential from B_z and the resulting horizontal components of the vector potential A_p^h (see Fisher *et al.*, 2010 for details). Thus, our A_p^h is determined in a staggered grid that is also exactly consistent with B_z .

The non-inductive component ψ is solved similarly as the \dot{P} potential in Equation 30 (including the *post facto* corrections) for each Assumption 1 and 2 (Section 2.3). When calculating the current density j_z for Assumption 2, the spatial derivatives of B_x and B_y were calculated using a central difference scheme, since we wish to keep j_z and thereby also ψ cospatial with the magnetogram data. However, since \mathbf{E}_I is defined on a staggered grid, the gradients $\nabla_x \psi$ and $\nabla_y \psi$ are not cospatial with E_x^1 and E_y^1 . This can be fixed by interpolating $\psi(x_m, y_n)$ to $\psi(x_{m+1/2}, y_{n+1/2})$, after which the gradients

$$\nabla_x \psi(x_m, y_{n+1/2}) = \frac{\psi(x_{m+1/2}, y_{n+1/2}) - \psi(x_{m-1/2}, y_{n+1/2})}{\Delta x} \tag{34}$$

$$\nabla_y \psi(x_{m+1/2}, y_n) = \frac{\psi(x_{m+1/2}, y_{n+1/2}) - \psi(x_{m+1/2}, y_{n-1/2})}{\Delta y} \tag{35}$$

are cospatial with E_x^1 and E_y^1 in Equations 31 and 32. Moreover, this approach ensures that $(\nabla \times \mathbf{E})_z$ in Faraday’s law is unaffected by $-\nabla_h \psi$.

Finally, it should be noted that the use of a staggered grid in the ELECTRICIT solutions of \mathbf{E} and A_p also requires (linear) interpolation back to a centered grid when Poynting or relative helicity fluxes are calculated.

Appendix C: Detection of Bad Pixels in our NOAA 11158 Series

When creating the NOAA 11158 magnetogram time series (Section 2.5), our algorithm detected at most 133 and on average 16 bad pixels in a single magnetogram cutout, which had an average shape of 745×630 pixels covering 4% of the solar disk. Ninety-nine percent of these pixels were detected in weak-field pixels where $B < 250 \text{ Mx cm}^{-2}$ (using the noise threshold chosen in Section 2.3). Maximum and average numbers of bad pixels in strong-field pixels were 6 and 0.2, bad pixels covering at most 0.03% and on average 7×10^{-6} of

the strong-field pixels. The number of bad pixels has a clear periodic behavior that follows the orbital motion of the SDO spacecraft and resulting variations in the formal error σ_B . The number of bad pixels detected in the weak-field region spikes consecutively with σ_B formal error, which is a direct consequence of using a fixed threshold for σ_B to detect the bad pixels. On the other hand, the number of bad pixels in the strong-field pixels spiked at the negative extremum of the spacecraft velocity when the σ_B was close to its minimum.

We have published the bitmaps of bad pixels and the corresponding B_{LOS} full-disk cutouts with the other relevant datasets of this paper at <https://zenodo.org/record/1034404>.

References

- Abbett, W.P., Mikić, Z., Linker, J.A., McTiernan, J.M., Magara, T., Fisher, G.H.: 2004, The photospheric boundary of Sun-to-Earth coupled models. *J. Atmos. Solar-Terr. Phys.* **66**, 1257. DOI. ADS.
- Aly, J.J.: 1991, How much energy can be stored in a three-dimensional force-free magnetic field? *Astrophys. J. Lett.* **375**, L61. DOI. ADS.
- Aschwanden, M.J., Sun, X., Liu, Y.: 2014, The magnetic field of Active Region 11158 during the 2011 February 12–17 flares: Differences between photospheric extrapolation and coronal forward-fitting methods. *Astrophys. J.* **785**, 34. DOI. ADS.
- Berger, M.A., Field, G.B.: 1984, The topological properties of magnetic helicity. *J. Fluid Mech.* **147**, 133. DOI. ADS.
- Bobra, M.G., Sun, X., Hoeksema, J.T., Turmon, M., Liu, Y., Hayashi, K., Barnes, G., Leka, K.D.: 2014, The Helioseismic and Magnetic Imager (HMI) vector magnetic field pipeline: SHARPs – Space-Weather HMI Active Region Patches. *Solar Phys.* **289**, 3549. DOI. ADS.
- Chen, P.F.: 2011, Coronal mass ejections: Models and their observational basis. *Living Rev. Solar Phys.* **8**, 1. DOI. ADS.
- Cheung, M.C.M., DeRosa, M.L.: 2012, A method for data-driven simulations of evolving solar active regions. *Astrophys. J.* **757**, 147. DOI. ADS.
- Cheung, M.C.M., De Pontieu, B., Tarbell, T.D., Fu, Y., Tian, H., Testa, P., Reeves, K.K., Martínez-Sykora, J., Boerner, P., Wülser, J.P., Lemen, J., Title, A.M., Hurlburt, N., Kleint, L., Kankelborg, C., Jaeggli, S., Golub, L., McKillop, S., Saar, S., Carlsson, M., Hansteen, V.: 2015, Homologous helical jets: Observations by IRIS, SDO, and Hinode and magnetic modeling with data-driven simulations. *Astrophys. J.* **801**, 83. DOI. ADS.
- DeForest, C.E., Hagenaar, H.J., Lamb, D.A., Parnell, C.E., Welsch, B.T.: 2007, Solar magnetic tracking. I. Software comparison and recommended practices. *Astrophys. J.* **666**, 576. DOI. ADS.
- Démoulin, P.: 2007, Recent theoretical and observational developments in magnetic helicity studies. *Adv. Space Res.* **39**(11), 1674.
- Fisher, G.H., Welsch, B.T.: 2008, FLCT: A fast, efficient method for performing local correlation tracking. In: Howe, R., Komm, R.W., Balasubramaniam, K.S., Petrie, G.J.D. (eds.) *Subsurface and Atmospheric Influences on Solar Activity*, *Astron. Soc. Pac. Conf. Ser.* **383**, 373. ADS.
- Fisher, G.H., Welsch, B.T., Abbett, W.P.: 2012, Can we determine electric fields and Poynting fluxes from vector magnetograms and Doppler measurements? *Solar Phys.* **277**, 153. DOI. ADS.
- Fisher, G.H., Welsch, B.T., Abbett, W.P., Bercik, D.J.: 2010, Estimating electric fields from vector magnetogram sequences. *Astrophys. J.* **715**, 242. DOI. ADS.
- Fisher, G.H., Abbett, W.P., Bercik, D.J., Kazachenko, M.D., Lynch, B.J., Welsch, B.T., Hoeksema, J.T., Hayashi, K., Liu, Y., Norton, A.A., Dalda, A.S., Sun, X., DeRosa, M.L., Cheung, M.C.M.: 2015, The coronal global evolutionary model: Using HMI vector magnetogram and Doppler data to model the buildup of free magnetic energy in the solar corona. *Space Weather* **13**, 369. DOI. ADS.
- Gibb, G.P.S., Mackay, D.H., Green, L.M., Meyer, K.A.: 2014, Simulating the formation of a sigmoidal flux rope in AR10977 from SOHO/MDI magnetograms. *Astrophys. J.* **782**, 71. DOI. ADS.
- Gosain, S., Pevtsov, A.A.: 2013, Resolving azimuth ambiguity using vertical nature of solar quiet-Sun magnetic fields. *Solar Phys.* **283**, 195. DOI. ADS.
- Hoeksema, J.T., Liu, Y., Hayashi, K., Sun, X., Schou, J., Couvidat, S., Norton, A., Bobra, M., Centeno, R., Leka, K.D., Barnes, G., Turmon, M.: 2014, The Helioseismic and Magnetic Imager (HMI) vector magnetic field pipeline: Overview and performance. *Solar Phys.* **289**, 3483. DOI. ADS.
- Inoue, S.: 2016, Magneto-hydrodynamics modeling of coronal magnetic field and solar eruptions based on the photospheric magnetic field. *Prog. Earth Planet. Sci.* **3**, 19. DOI. ADS.

- Jiang, C., Wu, S.T., Feng, X., Hu, Q.: 2016, Data-driven magnetohydrodynamic modelling of a flux-emerging active region leading to solar eruption. *Nat. Commun.* **7**, 11522. DOI. ADS.
- Jing, J., Park, S.-H., Liu, C., Lee, J., Wiegmann, T., Xu, Y., Deng, N., Wang, H.: 2012, Evolution of relative magnetic helicity and current helicity in NOAA Active Region 11158. *Astrophys. J. Lett.* **752**, L9. DOI. ADS.
- Jones, E., Oliphant, T., Peterson, P.: 2001, SciPy: Open source scientific tools for Python. <http://www.scipy.org/>.
- Kazachenko, M.D., Fisher, G.H., Welsch, B.T.: 2014, A comprehensive method of estimating electric fields from vector magnetic field and Doppler measurements. *Astrophys. J.* **795**, 17. DOI. ADS.
- Kazachenko, M.D., Fisher, G.H., Welsch, B.T., Liu, Y., Sun, X.: 2015, Photospheric electric fields and energy fluxes in the eruptive Active Region NOAA 11158. *Astrophys. J.* **811**, 16. DOI. ADS.
- Keller, C.U., Harvey, J.W., Giampapa, M.S.: 2003, SOLIS: An innovative suite of synoptic instruments. In: Keil, S.L., Avakyan, S.V. (eds.) *Innovative Telescopes and Instrumentation for Solar Astrophysics, Proc. SPIE 4853*, 194. DOI. ADS.
- Lagg, A., Lites, B., Harvey, J., Gosain, S., Centeno, R.: 2015, Measurements of photospheric and chromospheric magnetic fields. *Space Sci. Rev.* DOI. ADS.
- Leake, J.E., Linton, M.G., Schuck, P.W.: 2017, Testing the accuracy of data-driven MHD simulations of active region evolution. *Astrophys. J.* **838**(2), 113.
- Liu, Y., Schuck, P.W.: 2012, Magnetic energy and helicity in two emerging active regions in the Sun. *Astrophys. J.* **761**, 105. DOI. ADS.
- Liu, Y., Hoeksema, J.T., Sun, X., Hayashi, K.: 2017, Vector magnetic field synoptic charts from the Helioseismic and Magnetic Imager (HMI). *Solar Phys.* **292**, 29. DOI. ADS.
- Mackay, D.H., van Ballegoijen, A.A.: 2006, Models of the Large-Scale Corona. I. Formation, evolution, and liftoff of magnetic flux ropes. *Astrophys. J.* **641**, 577. DOI. ADS.
- Metcalf, T.R.: 1994, Resolving the 180-degree ambiguity in vector magnetic field measurements: The 'minimum' energy solution. *Solar Phys.* **155**, 235. DOI. ADS.
- Mickey, D.L., Canfield, R.C., Labonte, B.J., Leka, K.D., Waterson, M.F., Weber, H.M.: 1996, The imaging vector magnetograph at Haleakala. *Solar Phys.* **168**, 229. DOI. ADS.
- Mumford, S.J., Christe, S., Pérez-Suárez, D., Ireland, J., Shih, A.Y., Inglis, A.R., Liedtke, S., Hewett, R.J., Mayer, F., Hughtit, K., Freij, N., Meszaros, T., Bennett, S.M., Malocha, M., Evans, J., Agrawal, A., Leonard, A.J., Robitaille, T.P., Mampaey, B., Campos-Rozo, J.I., Kirk, M.S.: 2015, SunPy – Python for solar physics. *Comput. Sci. Discov.* **8**(1), 014009. DOI. ADS.
- Pagano, P., Mackay, D.H., Poedts, S.: 2013, Magnetohydrodynamic simulations of the ejection of a magnetic flux rope. *Astron. Astrophys.* **554**, A77. DOI. ADS.
- Pariat, E., Valori, G., Démoulin, P., Dalmasse, K.: 2015, Testing magnetic helicity conservation in a solar-like active event. *Astron. Astrophys.* **580**, A128. DOI. ADS.
- Pariat, E., Leake, J.E., Valori, G., Linton, M.G., Zuccarello, F.P., Dalmasse, K.: 2017, Relative magnetic helicity as a diagnostic of solar eruptivity. *Astron. Astrophys.* **601**, A125. DOI.
- Pesnell, W.D., Thompson, B.J., Chamberlin, P.C.: 2012, The Solar Dynamics Observatory (SDO). *Solar Phys.* **275**, 3. DOI. ADS.
- Pomoell, J., Lumme, E., Kilpua, E.K.J.: 2017, Modeling the eruption of coronal flux ropes using the time-dependent magnetofrictional method. *Solar Phys.* In preparation.
- Ravindra, B., Longcope, D.W., Abbott, W.P.: 2008, Inferring photospheric velocity fields using a combination of minimum energy fit, local correlation tracking, and Doppler velocity. *Astrophys. J.* **677**, 751. DOI. ADS.
- Riley, P., Lionello, R., Linker, J.A., Mikic, Z., Luhmann, J., Wijaya, J.: 2011, Global MHD modeling of the solar corona and inner heliosphere for the whole heliosphere interval. *Solar Phys.* **274**, 361. DOI. ADS.
- 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 Phys.* **275**, 207. DOI. ADS.
- Schrijver, C.J., Aulanier, G., Title, A.M., Pariat, E., Delannée, C.: 2011, The 2011 February 15 X2 flare, ribbons, coronal front, and mass ejection: Interpreting the three-dimensional views from the Solar Dynamics Observatory and STEREO guided by magnetohydrodynamic flux-rope modeling. *Astrophys. J.* **738**, 167. DOI. ADS.
- Schuck, P.W.: 2006, Tracking magnetic footpoints with the magnetic induction equation. *Astrophys. J.* **646**, 1358. DOI. ADS.
- Schuck, P.W.: 2008, Tracking vector magnetograms with the magnetic induction equation. *Astrophys. J.* **683**, 1134. DOI. ADS.
- Schuck, P.W., Antiochos, S.K., Leka, K.D., Barnes, G.: 2016, Achieving consistent Doppler measurements from SDO/HMI vector field inversions. *Astrophys. J.* **823**, 101. DOI. ADS.

- Seehafer, N.: 1978, Determination of constant alpha force-free solar magnetic fields from magnetograph data. *Solar Phys.* **58**, 215. DOI. ADS.
- Shibata, K., Magara, T.: 2011, Solar flares: Magnetohydrodynamic processes. *Living Rev. Solar Phys.* **8**, 6. DOI. ADS.
- Snodgrass, H.B.: 1983, Magnetic rotation of the solar photosphere. *Astrophys. J.* **270**, 288. DOI. ADS.
- Sun, X.: 2013, On the coordinate system of Space-Weather HMI Active Region Patches (SHARPs): A technical note. *ArXiv e-prints*. ADS.
- Sun, X., Hoeksema, J.T., Liu, Y., Wiegelmann, T., Hayashi, K., Chen, Q., Thalmann, J.: 2012, Evolution of magnetic field and energy in a major eruptive active region based on SDO/HMI observation. *Astrophys. J.* **748**, 77. DOI. ADS.
- Swarztrauber, P., Sweet, R.: 1975, *Efficient FORTRAN Subprograms for the Solution of Elliptic Partial Differential Equations*.
- Tremblay, B., Vincent, A.: 2015, A minimum energy fit method to reconstruct photospheric velocity and magnetic diffusivity in active regions from observed magnetograms and Dopplergrams. *Solar Phys.* **290**, 437. DOI. ADS.
- 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 Phys.* **249**, 167. DOI. ADS.
- Tziotziou, K., Georgoulis, M.K., Liu, Y.: 2013, Interpreting eruptive behavior in NOAA AR 11158 via the region's magnetic energy and relative-helicity budgets. *Astrophys. J.* **772**, 115. DOI. ADS.
- Tziotziou, K., Georgoulis, M.K., Raouafi, N.-E.: 2012, The magnetic energy-helicity diagram of solar active regions. *Astrophys. J. Lett.* **759**, L4. DOI. ADS.
- van Ballegoijen, A.A., Priest, E.R., Mackay, D.H.: 2000, Mean field model for the formation of filament channels on the Sun. *Astrophys. J.* **539**, 983. DOI. ADS.
- Weinzierl, M., Yeates, A.R., Mackay, D.H., Henney, C.J., Arge, C.N.: 2016, A new technique for the photospheric driving of non-potential solar coronal magnetic field simulations. *Astrophys. J.* **823**, 55. DOI. ADS.
- Welsch, B.T.: 2006, Magnetic flux cancellation and coronal magnetic energy. *Astrophys. J.* **638**, 1101. DOI. ADS.
- Welsch, B.T., Fisher, G.H., Sun, X.: 2013, A magnetic calibration of photospheric Doppler velocities. *Astrophys. J.* **765**, 98. DOI. ADS.
- Welsch, B.T., ABBETT, W.P., De Rosa, M.L., Fisher, G.H., Georgoulis, M.K., Kusano, K., Longcope, D.W., Ravindra, B., Schuck, P.W.: 2007, Tests and comparisons of velocity-inversion techniques. *Astrophys. J.* **670**, 1434. DOI. ADS.
- Welsch, B.T., Li, Y., Schuck, P.W., Fisher, G.H.: 2009, What is the relationship between photospheric flow fields and solar flares? *Astrophys. J.* **705**, 821. DOI. ADS.
- Wiegelmann, T., Petrie, G.J.D., Riley, P.: 2015, Coronal magnetic field models. *Space Sci. Rev.* DOI. ADS.
- Wiegelmann, T., Sakurai, T.: 2012, Solar force-free magnetic fields. *Living Rev. Solar Phys.* **9**, 5. DOI. ADS.
- Yang, W.H., Sturrock, P.A., Antiochos, S.K.: 1986, Force-free magnetic fields – The magneto-frictional method. *Astrophys. J.* **309**, 383. DOI. ADS.
- Yeates, A.R.: 2017, Sparse reconstruction of electric fields from radial magnetic data. *Astrophys. J.* **836**, 131. DOI. ADS.
- Yee, K.: 1966, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Trans. Antennas Propag.* **14**, 302. DOI. ADS.