An Evolution and Eruption of the Coronal Magnetic Field through a Data-driven MHD Simulation

Satoshi Inoue1, Keiji Hayashi2, and Takahiro Miyoshi3

1 Center for Solar-Terrestrial Research, New Jersey Institute of Technology, University Heights, Newark, NJ 07102-1982, USA; Satoshi.Inoue@njit.edu
2 College of Science, George Mason University, Fairfax, VA 22030, USA
3 Graduate School of Advanced Science and Engineering, Hiroshima University, 1-3-1 Kagamiyama, Higashihiroshima 739-8526, Japan

Received 2021 November 24; revised 2022 September 20; accepted 2022 October 13; published 2023 March 28

Abstract

We present a newly developed data-driven magnetohydrodynamics (MHD) simulation code under a zero-$\beta$ approximation based on a method proposed by Hayashi et al. 2018 and 2019. Although many data-driven MHD simulations have been developed and conducted, there are not many studies on how accurately those simulations can reproduce the phenomena observed in the solar corona. In this study, we investigated the performance of our data-driven simulation quantitatively using ground-truth data. The ground-truth data was produced by an MHD simulation in which the magnetic field is twisted by the sunspot motions. A magnetic flux rope (MFR) is created by the cancellation of the magnetic flux at the polarity inversion line due to the converging flow on the sunspot, which eventually leads to the eruption of the MFR. We attempted to reproduce these dynamics using the data-driven MHD simulation. The coronal magnetic fields are driven by the electric fields, which are obtained from a time series of the photospheric magnetic field that is extracted from the ground-truth data, on the surface. As a result, the data-driven simulation could capture the subsequent MHD processes, the twisted coronal magnetic field and formation of the MFR, and also its eruption. We report these results and compare them with the ground-truth data, and discuss how to improve the accuracy and optimize the numerical method.

Unified Astronomy Thesaurus concepts: Solar flares (1496); Magnetohydrodynamical simulations (1966); Solar active region magnetic fields (1975); Solar coronal mass ejections (310)

Supporting material: animations

1. Introduction

Active phenomena observed in the solar corona are widely believed to be driven by the coronal magnetic field (Shibata & Magara 2011). Thus, knowledge of how the coronal magnetic field evolves from the initial energy through its release is essential to understand these phenomena. Solar flares are one of the most common of these phenomena and are often linked to the origin site of coronal mass ejections (CMEs). CMEs bring a lot of coronal gas and magnetic field from the Sun to interplanetary space and cause massive electromagnetic disturbances in the Earth’s magnetosphere when they collide. Therefore, how the coronal magnetic field evolves not only influences the magnetic environment of the solar corona, but also extends this influence to near-Earth space and even to the furthest reaches of the heliosphere.

Magnetic field observations have been conducted by ground and space-based observations, which provide us with a time series of the photospheric magnetic field data in high temporal and spatial resolutions. For instance, the Solar Dynamics Observatory (Pesnell et al. 2012) provides the magnetic field every 12 minutes with a spatial resolution of 0’5, which is in Spaceweather HMI Active Patch format (Bobra et al. 2014). In the study in Wang et al. (2017), the Goode Solar Telescope (GST; Goode & Cao 2012) at the Bear Solar Observatory (BBSO) provides a magnetic field at 0’08 with 87 s cadence. However, the fact that the magnetic field is only measured at the photosphere is a technical problem, i.e., a three-dimensional (3D) magnetic field is not measured directly. A nonlinear force-free field extrapolation (NLFFF; e.g., Inoue 2016) is a useful tool that extrapolates a 3D magnetic field numerically based on the photospheric magnetic field under a force-free approximation. This technique offers several functions; for instance, time variation of the 3D magnetic structure before and after a flare (Schrijver et al. 2008), a formation process of the non-potential magnetic field leading to a solar flare (e.g., Inoue et al. 2011; Sun et al. 2012; Jiang et al. 2014; Woods et al. 2020), and analysis of solar flare onset in terms of time variations of the 3D magnetic fields (e.g., Muhamad et al. 2018; Kang et al. 2019; Yamasaki et al. 2021). However, NLFFF extrapolation only provides one snapshot of the magnetic field bounded by the force-free assumption and the essential problems still remain. For instance, the photospheric magnetic field used as the boundary condition cannot satisfy the force-free state (Metcalf et al. 1995; Kawabata et al. 2020a) and there is no guarantee of a unique solution under the given boundary and initial conditions (Kawabata et al. 2020b).

Although the NLFFF provides a snapshot of a 3D magnetic field at a specific time, a data-constrained magnetohydrodynamic (MHD) simulation that uses the potential field, NLFFF, and a non-force-free field as the initial conditions, can cover the evolution and dynamics of the magnetic field that is freed from the assumption of an equilibrium state (e.g., Inoue 2016; Muhamad et al. 2017; Inoue et al. 2018; Jiang et al. 2018; Prasad et al. 2020; Nayak et al. 2021). The data-constrained MHD simulation uses only one snapshot vector magnetic field. According to Inoue et al. (2018), the normal component of the photospheric magnetic field is fixed with
time and the time-dependent horizontal magnetic fields are derived from the induction equation, so the evolution of the tangential components is not exactly consistent with the time series of the photospheric magnetic field. Nevertheless, these simulations reproduce the observations quantitatively, such as the distribution of the flare ribbons and the direction of the eruption of the magnetic flux rope (MFR; e.g., Inoue et al. 2014; Jiang et al. 2017; Muhamad et al. 2017).

A data-driven MHD simulation drives the coronal magnetic field where a time series of the photospheric magnetic field is applied to the boundary condition (Toriiumi et al. 2020; Jiang et al. 2022). This method is more advanced than the data-constrained MHD simulation. The problem is how to treat the boundary condition, i.e., how to treat the induction equation at the bottom boundary. Note that we consider a very simple situation under zero-\(\beta\) approximation without taking into account the time-varying density and pressure at the bottom surface. The induction equation is written as follows in an ideal MHD framework:

\[
\frac{\partial B}{\partial t} = -\nabla \times E = \nabla \times (v \times B),
\]

where \(B, E\), and \(v\) are magnetic field, electric field, and plasma velocity, respectively. From Equation (1), we have several ways to drive the coronal magnetic field above; for instance, the electric field is given at the bottom boundary (Fisher et al. 2010, 2020; Kazachenko et al. 2014; Lumme et al. 2017; Hayashi et al. 2018). \(E\) has gauge invariance, i.e., it should be described as \(E = E_{\phi} + \nabla \phi\), which implies that the given electric field and the evolution of the coronal magnetic field depend on the gauge invariance. Nevertheless, Cheung & DeRosa (2012), Pomoell et al. (2019), and Price et al. (2020) reproduced various observational features by choosing the gauge with ingenuity. For example, Pomoell et al. (2019) compare the magnetic field lines with Extreme Ultraviolet (EUV) images and Price et al. (2020) superimpose the field lines on the images for a more rigorous comparison. The structure of the magnetic field lines is in good agreement with the structure of the magnetic field lines inferred from the images. Furthermore, Cheung & DeRosa (2012) reproduced the back reaction of the eruption that enhances the photospheric magnetic field after the eruption. Recently, Kaneko et al. (2021) derived the photospheric velocity by inversely solving the induction equation by imposing physical constraints on the gauge and driving the coronal magnetic field using the obtained velocity. Several methods directly provide either the magnetic field or velocity, or both in their data-driven simulation despite that both the magnetic field and velocity given at the boundary conditions is an over-specification for an MHD. For example, the velocity is derived from the algorithm Differential Affine Velocity Estimator for Vector Magnetogram (DAVE4VM; Schuck 2008). Although the detailed method is different for each, they produce a non-potential magnetic field prior to the flare or the eruption of the MFR well (Jiang et al. 2016, 2021; Leake et al. 2017; Guo et al. 2019; Hayashi et al. 2019; Liu et al. 2019).

Recently, Hayashi et al. (2018, 2019) proposed a new method for the data-driven MHD simulation in which three electric fields are defined at the bottom boundary, and a half grid above and below that completely reproduces the photospheric magnetic field through the induction equation. Hayashi et al. (2019) include a data-driven simulation in which the coronal magnetic field is driven by the velocity. Following their method, the velocity, which is obtained from DAVE4VM, is given on the bottom surface while the magnetic field is driven by the velocity through the MHD process. Their simulation was applied to solar active region (AR) 11158 and reproduced the evolution toward the non-potential magnetic field starting with the potential field. Eventually, the magnetic fields prior to the flare were obtained, which were similar to the coronal magnetic fields inferred from the EUV observations. In this study, we develop a data-driven MHD code based on the method proposed by Hayashi et al. (2018). The code is designed with a zero-\(\beta\) approximation because the magnetic pressure is dominant over the gas pressure in the lower solar corona, and constructed using a finite-differential method based on that in Inoue et al. (2014). The purpose of this study is to test the accuracy of this data-driven MHD simulation to compare with the ground-truth data, including the evolution of the coronal magnetic fields during the energy buildup and eruption process. Leake et al. (2017) and Toriumi et al. (2020) discussed the accuracy of the data-driven simulations by comparison to the ground-truth data. Leake et al. (2017) examined the sensitivity of data-driven simulation results to the cadence of the input boundary driving-data maps. In their study, all MHD variables are given at the bottom boundary. They found that the cadence is indeed important for reproducibility. Toriumi et al. (2020) compare the simulation results from the participating data-driven simulation models. In their comparison, the models are given only the information on the boundary magnetic field at a low cadence and yield very different solutions from the ground-truth solution. Our electric-field-driven model uses only the information on the solar-surface magnetic field. Therefore, our newly developed code also needs an accuracy test.

From the two aforementioned works, in this study, the target ground-truth three-dimensional solution and the bottom-boundary time-dependent magnetic field data maps are prepared with the same MHD code but driven with variables other than the magnetic field. We followed the velocity-driven MHD model procedures in Amari et al. (2000, 2003) to generate the ground-truth solutions and the time-dependent magnetic-field data maps. This ground-truth data includes the eruption of the magnetic field, which is different from previous works by Leake et al. (2017) and Toriumi et al. (2020).

Note that the electric field obtained from the inversion of the induction equation is not uniquely determined due to the uncertainty of the gauge; hence it is not guaranteed that the electric field will indeed reproduce the evolution of the magnetic field in the ground-truth data produced by velocity-driven MHD. In addition, the electric-field inversion is designed to utilize only (temporal difference of) the magnetic field data on the bottom-boundary surface at two consecutive sampling instants. In other words, the method in Hayashi et al. (2018) is not designed to accommodate information on the temporal evolution of the ground-truth magnetic field or the three-dimensional structure of the whole volume. Given these caveats, it is crucially important to assess how well electric-field-driven MHD models using the electric-field inversion can yield the ground-truth solution. In this work, we quantitatively evaluate the results of the data-driven MHD simulation by comparing them with the ground-truth data. We further discuss
the numerical treatment of the implementation method, for which intrinsically we are allowed to make a somewhat arbitrary choice in, for example, spatial differencing.

The rest of this paper is organized as follows. The numerical method is described in Section 2. The results and discussion are presented in Sections 3 and 4. Finally, our conclusion are summarized in Section 5.

2. Numerical Method

2.1. Basic Equations

We solve the following zero-$\beta$ MHD equation to create the ground-truth data and also conduct the data-driven simulation,

$$\frac{\partial p}{\partial t} = -\nabla \cdot (\rho v) + \zeta \nabla^2 \rho, \quad (2)$$

$$\frac{\partial v}{\partial t} = -(v \cdot \nabla)v + \frac{1}{\rho} J \times B + \nu \nabla^2 v, \quad (3)$$

$$\frac{\partial B}{\partial t} = \nabla \times (v \times B) + \eta \nabla^2 B - \nabla \phi, \quad (4)$$

$$J = \nabla \times B, \quad (5)$$

$$\frac{\partial \phi}{\partial t} + c_s^2 \nabla \cdot B = -\frac{c_s^2}{\rho} \phi, \quad (6)$$

where $B$ is the magnetic flux density, $v$ is the velocity, $J$ is the electric current density, $\rho$ is the pseudo density, and $\phi$ is the convenient potential to remove errors derived from $\nabla \cdot B$ (Dedner et al. 2002), respectively. $\rho_0$ corresponds to the initial density. The length, magnetic field, density, velocity, time, and electric current density are normalized by $L^*, B^*, \rho^*, \nu \equiv B^*/(\mu_\circ \rho_0)^{1/2}$, where $\mu_\circ$ is the magnetic permeability, $\nu^* = L^*/V_\Lambda^*$, and $J^* = B^*/\mu_\circ L^*$. $\nu$ and $\eta$ are viscosity and resistivity, fixed by $1.0 \times 10^{-3}$ and $1.0 \times 10^{-5}$, respectively, and the coefficients $c_s^2, c_s^2'$ in Equation (6) also fix the constant value, 0.04 and 0.1, respectively. $\zeta$ is the diffusion coefficient of the density, which avoids the sudden variation of the density to improve the robustness of the simulation, where $\zeta$ is set to $1.0 \times 10^{-4}$ in this study. A numerical box of $1.0 \times 1.0 \times 1.0$, which is given in its non-dimensional value, is divided by $320 \times 320 \times 320$ grid points.

2.2. Ground-truth Data

We first establish the ground-truth data according to an MHD simulation done by Amari et al. (2000, 2003) to verify the reproducibility of the energy storage and release processes of the coronal magnetic field by our data-driven MHD simulation. We first set a simple dipole magnetic field like sunspots as shown in Figure 1(a), from which the potential field is extrapolated in 3D space following the Green function method (Sakurai 1982) as shown in Figure 1(b). The initial velocity and density are set as $|v| = 0.0$ and $\rho_0 = 1.0$ in the whole space, respectively.

Next, we impose a twisting motion on the photosphere according to

$$v_{\psi}^{(b)}|_{t=0} = (v_x, v_y) = \left(\frac{\partial \psi}{\partial y}, -\frac{\partial \psi}{\partial x}\right), \quad (7)$$

where $\psi$ is a stream function that satisfies $\bar{v}_B^{(b)}|_{t=0} = z \times \nabla h \psi$. We give $\psi$ the following formula:

$$\psi(x, y, t) = \gamma(t) \{ B_z^{(b)} \} \epsilon \left( \frac{e^{-\frac{(x-x_0)^2}{\sigma_x^2}}}{\sigma_x^2} \right),$$

and

$$\gamma(t) = -0.5 \tanh \left\{ \frac{2.0 (t - t_{cri}) - 1.0}{0.5} \right\} + 0.5,$$

where $B_z^{(b)}$ and $B_z^{(b)}_{\max}$ correspond to the magnetic field measured at the photosphere and the maximum value, respectively. If time $t$ is beyond the critical time set at $t_{cri}$ where we set $t_{cri} = 9.0$, $\gamma(t)$ quickly falls to zero. The twisting motion is shown in Figure 1(c). It is convenient to give $\psi$ as a function of $B_z$, because the twisting motion is imposed along the contour of $B_z$, i.e., $(v_x, \nabla \cdot B)_{B_z=0}$ is promised regardless of $B_z$. Thus, the distribution of $B_z$ is not be changed by the twisting motion. Since we assume $v_z = 0$ at the bottom surface as well as $(v_x, \nabla \cdot B)_{B_z=0}$, these conditions satisfy $\partial B_z = 0$. Therefore, normal flux is not transported across the bottom surface during the evolution and our simulation can keep $|B_z| dS = 0$. When $\gamma(t)$ falls to zero, the twisting motion at the photosphere comes to a complete halt. Afterward, the magnetic field is relaxed for a while, i.e., no external motion is imposed. The horizontal magnetic components $B_x$ and $B_y$ at the photosphere are following the equations during the twisting motion and relaxation process after the twisting motion is over,

$$\frac{\partial B_x^{(b)}}{\partial t} = -\left( \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} \right) - \frac{\partial \phi}{\partial x}, \quad (8)$$

$$\frac{\partial B_y^{(b)}}{\partial t} = -\left( \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} \right) - \frac{\partial \phi}{\partial y}, \quad (8)$$

while $B_z$ is fixed. $E_x$ and $E_y$ represent the electric field in the $x$ and $y$ components, which are described as $-(v_x \times B) \cdot x$ and $-(v_y \times B) \cdot y$, respectively. The $v_x$ and density $\rho$ are fixed with the initial value at the bottom boundary and $\phi$ is imposed on the Neumann condition throughout the simulation. Note that, during the relaxation process, $E_z$ at the bottom surface is zero, i.e., $\partial E_z = \partial_x E_z = 0$ because the evolutions of $v_x$ and $v_y$ are halted there ($v_x = v_y = 0$).

Finally, after the relaxation, we impose a diverging motion on the sunspot, which corresponds to a converging motion around the polarity inversion line (PIL), as shown in Figure 1(d). The diverging motion is given as the following formula (e.g., Xia et al. 2014):

$$v_x^{(d)} = -\frac{\partial [B_x^{(d)}]}{\partial x} e^{-\frac{(x-x_0)^2}{\sigma_x^2}},$$

$$v_y^{(d)} = -\frac{\partial [B_y^{(d)}]}{\partial y} e^{-\frac{(y-y_0)^2}{\sigma_y^2}}, \quad (9)$$

where $x_d = y_d = 0.1$ is given. The magnetic field at the bottom boundary follows an induction equation in each component as
Note that the diffusion is given in the direction of converging motion (in this study, it corresponds to the $x$ direction) to make the simulation more robust. This diverging motion is continuously imposed on the sunspot until the end of the simulation. In both cases, twisting motion and diverging motion, we normalized the velocity $v^{(b)}$ by the maximum value $|v_{y:}\text{max}|$ and multiplied it by 0.01, i.e., $v^{(b)} \Rightarrow 0.01v^{(b)}/|v_{y:}\text{max}|$, before being used in Equations (8) and (10), so that the maximum absolute value of the velocity is set as 0.01 at the bottom surface.

### 2.3. Data-driven Simulation

#### 2.3.1. Three Electric Fields Driving the Coronal Magnetic Field

In this study, a data-driven MHD simulation is performed according to the method proposed by Hayashi et al. (2018, 2019). We now briefly describe this method. In this method, the coronal magnetic field is driven by electric fields given at the bottom. In order to determine the electric fields that derive the time-evolving observed photospheric magnetic field through an induction equation, three types of electric fields are assumed to satisfy the following induction equations:

$$\frac{\partial B_z}{\partial t} = -\mathbf{z} \cdot \nabla \times \mathbf{E}^{(1)},$$

Figure 1. (a) A bipolar magnetic field given at the bottom surface where red and blue correspond to the positive and negative polarities, respectively. The full simulation region is plotted in the $X$- and $Y$-directions. (b) The potential magnetic field that is extrapolated from the bipolar magnetic field is shown in (a). This is used as the initial condition of the MHD simulation. (c) Twisting motion of the velocity given on the bipolar field, which can maintain the distribution of the bipolar field. (d) Diverging motion of the velocity given on the bipolar field. This velocity plays a role in converging motion at the PIL.
\[
\frac{\partial B_{h,df}}{\partial t} = -\nabla \times E^{(2)},
\]
(12)
\[
\frac{\partial B_{h,cf}}{\partial t} = -\nabla \times E^{(3)},
\]
(13)
where the horizontal field \((B_h)\) at the photosphere is decomposed into \(B_{h,df}\) and \(B_{h,cf}\), i.e., \(B_h = B_{h,df} + B_{h,cf}\), where these satisfy \(\nabla_h \times B_{h,df} = 0\) and \(\nabla_h \times B_{h,cf} = 0\), respectively.

Following Hayashi et al. (2018), three Poisson equations are derived from the above three induction equations. The first Poisson equation is obtained as follows:
\[
\nabla^2 \Phi^{(1)} = -\frac{\partial B_z}{\partial t},
\]
(14)

based on Equation (11) where we make the following assumption, \(E^{(1)} = z \times \nabla \Phi^{(1)}, \) i.e., \(\partial E_z^{(1)} / \partial z = 0\). Since the right-hand side of Equation (14) can be calculated from the photospheric magnetic field, we can solve Equation (14) under an appropriate boundary condition and obtain \(E^{(1)}\) to determine \(B_z\) at the photosphere.

The second Poisson equation is
\[
\nabla^2 \Phi^{(2)} = -z \cdot \left( \nabla \times \frac{\partial B_{h,df}}{\partial t} \right)
= -z \cdot \left( \nabla \times \frac{\partial B_h}{\partial t} \right),
\]
(15)

where \(E^{(2)} = -z \Phi^{(2)}\) is assumed. The advantage of this assumption is that it does not change the normal component of the magnetic field at the photosphere because it is determined by \(E_x^{(1)}\) and \(E_y^{(1)}\). Namely, if we further give \(E_x^{(2)}\) and \(E_y^{(2)}\), \(B_z\) changes and deviates from the observation. This assumption is able to avoid the conflict of \(B_z\) being obtained from \(E^{(1)}\) and \(E^{(2)}\). A detailed derivation process of Equation (15) is described in Appendix A.1.

The third Poisson equation is
\[
\nabla^2 \Phi^{(3)} = \frac{1}{2} \Delta z \nabla_h \cdot \left( \frac{\partial B_{h,cf}}{\partial t} \right)
= \frac{1}{2} \Delta z \nabla_h \cdot \left( \frac{\partial B_h}{\partial t} \right),
\]
(16)

where \(\Delta z\) is the grid interval and we make the following assumption: \(E^{(3)}_{|z=+\Delta z/2} = z \times \nabla_h \Phi^{(3)}\). Note that since \(B_z\) is perfectly derived from \(E_x^{(1)}\) and \(E_y^{(1)}\), \(E^{(3)}\) is an over condition to determine \(B_z\). Therefore, \(E^{(3)}\) is defined at the locations a half grid above and below the photosphere as follows:
\[
E^{(3)}_{|z=+\Delta z/2} = -E^{(3)}_{|z=-\Delta z/2}.
\]
(17)
\(\Phi^{(3)}\) is defined at the plane half grid above the bottom surface but it is determined by the given horizontal magnetic field at the bottom surface, so \(\Delta z/2\) plays the role of a bridge between the left and right-handed values defined at each different height. A detailed derivation process of Equation (16) is described in Appendix A.2. Eventually, we obtain \(E^{(1)}\), \(E^{(2)}\), and \(E^{(3)}\) through the three Poisson Equations (14)–(16). The positional relationship of the three electric fields is summarized in Figure 2.

### 2.3.2. How the Magnetic Field at the Bottom Surface is Driven

The magnetic field at the bottom surface is driven by \(E^{(1)} = (E_x^{(1)}, E_y^{(1)}, 0)\), \(E^{(2)} = (0, 0, E_z^{(2)})\), and \(E^{(3)} = (E_x^{(3)}, E_y^{(3)}, 0)\), where \(E^{(1)}\) and \(E^{(2)}\) are defined at the bottom surface while \(E^{(3)}\) is defined one half grid above (and also below) the bottom surface. Note that, following Hayashi et al. (2018), \(E^{(1)}\) is also placed at the surface one half grid above (and also below) the bottom, taking into account the conservation of the normal flux. The difference from \(E^{(3)}\) is that we assume \(E_z^{(1)} = E_z^{(3)} = -\Delta z/2\), so its \(z\)-derivative becomes zero. This positional relationship is represented in Figure 2.

The time derivative of \(B_z\) is written as
\[
\frac{\partial B_z}{\partial t} = -\left( \frac{\partial E_x^{(1)}}{\partial x} - \frac{\partial E_y^{(1)}}{\partial y} \right).
\]
(18)

and since \(E^{(2)}\) only has a \(z\) component, \(B_{h,cf}\) is described as
\[
\frac{\partial B_{h,cf}}{\partial t} = -\left( \frac{\partial E_z^{(2)}}{\partial y} - \frac{\partial E_z^{(2)}}{\partial x} \right),
\]
(19)

where all the components are defined in \(x\) and \(y\) components; thus, we omit the \(z\) component. Regarding the time derivative of \(B_{h,cf}\) by taking into account the assumption of \(E^{(3)}\), \(E_z^{(3)}\) equals zero and \(E^{(3)}\) is defined as the plane half grid above and below the bottom surface. The \(x\) component of \(\partial E_{h,cf}^{(1)}\) is written as
\[
\frac{\partial E_x^{(3)}}{\partial t} = \frac{\partial E_x^{(3)}}{\partial z}
\]

\[
E^{(3)}_{|z=+\Delta z/2} - E^{(3)}_{|z=-\Delta z/2} = \frac{2}{\Delta z} E^{(3)}_x,
\]

where the details are described in the Appendix of Hayashi et al. (2018). Therefore,
\[
\frac{\partial B_x}{\partial t} = -\left( \frac{\partial E_x^{(2)}}{\partial y} - \frac{2}{\Delta z} E^{(3)}_x \right),
\]

\[
\frac{\partial B_y}{\partial t} = -\left( \frac{\partial E_y^{(2)}}{\partial x} + \frac{2}{\Delta z} E^{(3)}_x \right),
\]

\[
\frac{\partial B_z}{\partial t} = -\left( \frac{\partial E_z^{(1)}}{\partial x} - \frac{\partial E_z^{(1)}}{\partial y} \right)
\]

are given at the bottom surface.

### 2.4. Numerical Method of the Data-driven Simulation

The same MHD equations (Equations (2)–(6)) are applied to the data-driven MHD simulation while the electric fields, \(E^{(1)}–(3)\) are given at the bottom and a half grid above and below as shown in Figure 2. Since the location that is one grid above
the bottom \((k = 1\) shown in Figure 2) corresponds to the practical bottom boundary, the derivation of the magnetic field at this position \((k = 1\) is important for the data-driven simulation in this study. Especially, the derivative in the \(z\) direction introduced in the induction equation should be handled with care because it includes the difference between non-MHD components \((E^{(1)}\) and \(E^{(3)}\), which are placed at \(k = 0\) or \(1/2\) and the electric fields obtained from MHD process \((-v \times B)\), which are placed at \(k = 2\).

According to a second-order finite-differential method, the derivative of \(E\) in the \(z\) direction at \(k = 1\) in the induction equation is written as

\[
\frac{\partial E}{\partial z}|_{k=1} = \frac{E_{k=2} - E_{k=0}}{2\Delta z},
\]

\[
= \frac{E_{k=2} - E_{k=1}}{2\Delta z} + \frac{E_{k=1} - E_{k=0}}{2\Delta z},
\]

\[
= \frac{E_{k=2} - E_{k=1}}{2\Delta z} + \left( \frac{E_{k=1} + E_{k=0}}{2\Delta z} \right) - \frac{2E_{k=0}}{2\Delta z},
\]

\[
= \frac{E_{k=2} - E_{k=1}}{2\Delta z} + \left( \frac{E_{k=1} + E_{k=0}}{2\Delta z} \right) - \frac{2E_{k=0}}{2\Delta z},
\]

where \(E\) is either the \(x\) or \(y\) component and we use the following approximation,

\[
E_{k=\frac{1}{2}} = \frac{E_{k=1} + E_{k=0}}{2}
\]

Figure 2. The positions of each electric field, \(E^{(1)}, E^{(2)}, \) and \(E^{(3)}\) that drive the coronal magnetic field, proposed by Hayashi et al. (2018), \(z = 0\) corresponds to the bottom boundary (photosphere) and \(z = \Delta z\) corresponds to a location one grid above the bottom surface where \(\Delta z\) is the distance between the neighboring grids set in this study. All physical values are defined by the black circles in this simulation where \(k\) takes an integer value. \(E^1\) and \(E^3\) are given at \(k = \pm 1/2\). The right square provides a summary of the electric fields that are given at \(k = 0\) and \(k = \pm 1/2\).

\[
E^{(3)} = (E_x^{(3)}, E_y^{(3)}, 0)
\]

\[
E^{(1)} = (E_x^{(1)}, E_y^{(1)}, 0)
\]

\[
E^{(2)} = (0, 0, E_z^{(2)})
\]

\[
E^{(3)} = -E^{(3)}
\]

\[
E^{(1)} = (E_x^{(1)}, E_y^{(1)}, 0)
\]

to obtain the last equation. When we take into account \(E_{k=\frac{1}{2}} = E^{(1)} + E^{(3)}\) and \(E_{k=0} = E^{(1)}\), we obtain

\[
\frac{\partial E}{\partial z}|_{k=1} = \frac{E_{k=2} - E_{k=1}}{2\Delta z} + \frac{E^{(3)}}{\Delta z}.
\]  

(22)

On the other hand, we can write the different formula as follows:

\[
\frac{\partial E}{\partial z}|_{k=1} = \frac{E_{k=2} - E_{k=0}}{2\Delta z},
\]

\[
= \frac{E_{k=2} + E_{k=1}}{2\Delta z} - \frac{E_{k=1} + E_{k=0}}{2\Delta z},
\]

\[
= \frac{E_{k=2} + E_{k=1}}{2\Delta z} - \frac{E_{k=1}}{\Delta z},
\]

where \(E_{k=\frac{1}{2}}\) is approximated in Equation (21). Eventually, we obtain the following equation:

\[
\frac{\partial E}{\partial z}|_{k=1} = \frac{E_{k=2} + E_{k=1}}{2\Delta z} - \frac{E^{(1)} + E^{(3)}}{\Delta z}.
\]

(23)

Hereafter, the former written in Equation (22) is denoted as Type A and the latter (Equation (23)) is denoted as Type B. We will mainly show the results obtained from Type A and discuss the difference in the results between Type A and Type B.

The coronal magnetic fields are driven as outlined in Figure 3. We derive the electric field \(E^{(1)-(3)}\) at each time interval from the two output data at \(t_n\) and \(t_{n+1}\) \((n = 1, 2, 3\cdots)\) obtained from the referenced MHD simulation where only the magnetic field at the bottom is used. The time interval \(t_{n+1} - t_n\)
is set to 0.5625. The electric fields are derived from the three Poisson Equations \((14) - (16)\). In this study, we use the Gauss–Seidel method to solve them. The performance and convergence of the Poisson equations are shown in the Appendix (see Figure A1). In the data-driven simulation, each electric field drives the upper coronal magnetic field at each interval. Note that, in this study, we drive the coronal magnetic field when the magnetic field \(B_{\text{xyz}} = B_x + B_y + B_z\) measured at the bottom is more than 0.0625, i.e., the weak magnetic field region does not change with time.

As can be seen in Table 1, we ran eight different simulations, which are mainly classified into the cases of TW and the cases of CV where TW and CV mean data-driven simulations in the twisting phase (energy buildup phase: \(t = 0 \sim 11.25\) in the MHD simulation) and the converging phase (energy released phase: \(t \geq 22.5\) in the MHD simulation), respectively. The initial conditions of the data-driven MHD simulations in the cases of TW and CV employ the magnetic fields at \(t = 0\) and \(t = 22.5\) obtained from the ground-truth data. Note that the data-driven simulations are not conducted continuously from \(t = 0\) to after the eruption because one of the objectives of this study is to determine how the data-driven simulation can produce the energy release phase. If we conduct the data-driven simulation from \(t = 0\) to the eruption, it becomes difficult to determine whether the errors are due to the energy buildup phase \((t = 0 \sim 11.5)\) or the relaxation phase \((t = 11.5 \sim 22.5)\). We focus on evaluating the data-driven simulation in the energy buildup phase in this study.
buildup phase without the extra errors. The Type A and B differential methods denote A ($A^0, A^1,$ and $A^2$), and B, respectively. Type A is further classified into $A^0, A^1,$ and $A^2$ where the vertical velocity ($v_z$) placed at one grid ($k=1$) above the bottom surface has been given through a specific update in the temporal evolution and the method is different in each case. In the cases of $A^1$ and $A^2$, the vertical velocity ($v_z$) is fixed to zero or is updated through a linear interpolation with values of $k=0$ and $k=2$, respectively, while $A^0$ has no specific update, i.e., the vertical velocity is derived from an equation of motion directly. In case B, we run one simulation where the vertical velocity at $k=1$ is updated using linear interpolation.

3. Results

3.1. Overview of Ground-truth Data

The temporal evolution of the magnetic and kinetic energies is shown in Figure 4. Both energies are built up by $t \sim 11.25$ due to the twisting motion of the sunspot, then gradually decrease by $t \sim 22.5$ in the relaxation phase. Note that the twisting motion is halted at $t \sim 11.25$ and the magnetic fields are relaxed by $t \sim 22.5$. Within a few moments of turning on the converging motion at $t \sim 22.5$, the kinetic energy increases dramatically, while the magnetic energy decreases, a lot of which is dissipated and converted into the Joule heating through the flux cancellation.

The evolution of the 3D magnetic field lines, while the twisting motion is imposed, is shown in Figure 5. As time passes, the magnetic field lines are sheared gradually and eventually a sigmoidal structure is formed at $t = 10.12$, which is often observed prior to flares (e.g., Inoue et al. 2012; Savcheva et al. 2012; Kawabata et al. 2018).

After the relaxation, we impose the diverging flow on the sunspots as shown in Figures 6(a)–(c) with the evolution of the sunspots. From the upper panels, a part of the magnetic flux is transported toward the PIL at which they are canceled. The
lower panels (Figures 6(d)–(e)) show the evolution of the magnetic field lines. The footpoints of the magnetic field lines are carried by the flow on both the sunspots and field lines of opposite polarity encounter each other at the PIL. Magnetic reconnection then takes place, resulting in the formation of the long and highly twisted lines as can be seen in Figures 6(e) and (f). The continuous converging flow around the PIL enhances the further twisted MFR and lifts it upward as can be seen in Figure 7(b) where the magnetic field lines just before the diverging flow are imposed and are shown in Figure 7(a). The current sheet is formed under the lifted MFR and eventually it erupts as shown in Figures 7(c) and (d).

We have attempted to reproduce these MHD processes by using the data-driven MHD simulation in which the electric field works as the driving source on the bottom surface, instead of the velocity field. These electric fields are derived from a time series of the magnetic field on the bottom surface of the ground-truth data produced by the MHD simulation.

3.2. Results of the Data-driven MHD Simulation

3.2.1. Energy Buildup Phase

First, we show the results of data-driven simulation in the energy buildup phase in which the ground-truth data shows the formation of a sigmoid. The initial condition of the data-driven simulation was given by the potential magnetic field shown in Figure 1(b) and the time-dependent electric fields are given at the bottom and one-half at the top (and also below). Figures 8(a) and (b) show the temporal evolution of the magnetic and kinetic energies, respectively. From Figure 8(a), we can see that the temporal evolutions of magnetic energies in the data-driven simulations are almost indistinguishable from each other. It should also be noted that the quantities from the data-driven simulations are very similar to the initial ground-truth data, but as the evolution continues as they deviate from this. Note that free magnetic energy is essential to discuss the eruption rather than the net magnetic energy shown in this study. However, the potential field is exactly the same in the ground-truth and data-driven simulations because the bottom $B_z$ is exactly the same between them. Therefore, the profile of the free magnetic energy is the same as the one shown in Figure 5(a) while the magnitudes differ between the net magnetic energy and free magnetic energy. From Figure 8(b), although the kinetic energy increases in the temporal evolution in both the ground-truth data and each data-driven simulation, the behaviors are slightly different among them. We note that the kinetic energy is much smaller than the magnetic energy.

The temporal evolution of the 3D magnetic field obtained from the data-driven simulation (TW-A02) is shown in Figure 9. The data-driven simulation reproduces an evolution similar to the ground-truth data as shown in Figure 5. In particular, the sigmoidal structure is reproduced well in the final time step. On the other hand, the sheared field lines accumulated in the data-driven simulation look weaker than those in the ground-truth data. Nevertheless, this result shows that the data-driven simulation works well.
We show more quantitative results by using magnetic twist defined as

$$T_w = \int \frac{\nabla \times \mathbf{B} \cdot \mathbf{B}}{|\mathbf{B}|^2} dl,$$  

(24)

where $dl$ is a line element of each field line (Berger & Prior 2006). Figure 10(a) shows the temporal evolution of the magnetic flux in the ground-truth data and each case of the data-driven simulation. These magnetic fluxes are stored by the magnetic field lines, which satisfy the condition $T_w \leq -0.1$. Although the magnetic flux, which is composed of the non-potential component, obtained from the data-driven simulation is a bit smaller than the magnetic flux obtained from the MHD simulation, the behavior is almost the same among them. Figure 10(b) shows a histogram of the magnetic flux, which depends on the twist value for the ground truth and data-driven simulation (TW-A02). Although the twist value obtained from the data-driven simulation is slightly weaker than that calculated from the ground-truth data, the two distributions are very similar. The distributions of the magnetic twist of each field line are mapped on each surface. The results calculated from the ground-truth data and the data-driven MHD (TW-A02) simulation are also very similar as can be seen in Figures 10(c) and (d). Thus, these results support that the data-
Figure 8. (a) Temporal evolution of the magnetic energy in the energy buildup phase where the black, purple, light blue, and red lines represent the results in the ground-truth data and each case of the data-driven MHD simulation, TW-A0, TW-A1, and TW-A2, respectively. These lines obtained from the data-driven simulations are almost overlapping. (b) Temporal evolution of the kinetic energy. The format is the same as in (a).

Figure 9. (a)–(d) Temporal evolution of the 3D magnetic field lines obtained from the data-driven MHD simulation (TW-A2). The format of this figure is the same as the one in Figure 5. An animation of the temporal evolution is available in the online Journal. The animation proceeds from $t = 0$ to 11.25, showing the formation of the S-shaped structure in the data-driven simulation. (An animation of this figure is available.)
driven simulation reproduces the twisting process of the magnetic field lines well.

3.2.2. Energy Release Phase

Next, we show the results obtained from the data-driven MHD simulation in the energy release phase. This corresponds to the eruption phase of the MFR formed by the converging motion. The initial condition of the data-driven simulation was given by the magnetic field at \( t = 22.5 \) of the ground-truth data. Figures 11(a) and (b) show the temporal evolutions of the magnetic and kinetic energies obtained from the ground-truth and data-driven simulations. We can see that the magnetic and kinetic energies obtained from the data-driven simulations capture the tendency to decrease and increase as is seen in the respective ground-truth data. However, in the case CV-A0 the kinetic energy grows very slowly compared to other cases, which will be discussed later.

As can be seen in Figure A3, the bottom \( B_z \) obtained from the data-driven simulation reproduces the ground-truth data well. The correlation coefficient is over 0.99 at each time. This means that the potential fields extrapolated from the bottom \( B_z \) of the ground-truth data and obtained from the data-driven simulation are almost identical. Although the magnitudes of the net magnetic energy and free magnetic energy are different from each other, the behavior is the same as that shown in Figure 11(a).

The temporal evolution of the 3D magnetic field lines in the case of CV-A2 is shown in Figure 12. The MFR is formed through the converging motion, which is driven by the electric field, and causes an eruption, under which the current sheet is enhanced. We confirmed that the evolving 3D magnetic field shown in the data-driven simulation is...
consistent with the ground-truth data; thus, the data-driven simulation appears to work well in the energy release process of the magnetic field that corresponds to the eruption phase of the MFR.

The results of the quantitative analysis (especially focusing on the case of CV-A2) are shown in Figure 13. Figure 13(a) shows the temporal evolution of the magnetic flux, which is dominated by the highly twisted field lines, $|T_{\nu}| \geq 1.0$, in each case where these highly twisted field lines are newly created during the evolution. The evolution obtained from the data-driven simulation in each case almost captures the ground-truth data (black) while the growth in the case of CV-A0 is very slow as inferred from the result shown in Figure 11(b). Figure 13(b) represents a histogram of the magnetic flux versus magnetic twist for the ground-truth data and data-driven simulation (CV-A2). Figures 13(c) and (d) show the distribution of the $T_{\nu}$ mapped on the surface, obtained from the ground-truth data and the data-driven simulation (CV-A2), respectively. Although these do not match exactly, both are very similar. Thus, we confirmed that the data-driven simulation works well in terms of the quantitative analysis.

We trace the magnetic axis of the MFR in the evolution for the ground-truth data and the data-driven MHD simulations, respectively. In this study, we detect the axis at which the sign of $B_z$ inverts along the center of the numerical box, which corresponds to the dashed vertical line as shown in Figure 14(a) because the symmetry of the MFR is good during the evolution. Figure 14(b) plots the field lines on the vertical cross section ($x$–$z$ plane) and the red circle points out the same position shown in Figure 14(a), which corresponds to the center of the MFR. The temporal evolutions of the MFR axis of the ground-truth data and those obtained from each case of the data-driven simulation are shown in Figure 14(c). Although case CV-A0 deviates from the result from the ground-truth data, other cases obtained from the data-driven simulations capture its behavior well. Therefore, the data-driven simulation developed in this study, which is based on that in Hayashi et al. (2018), can be used as a powerful tool to understand the evolution of the coronal magnetic field and the physics of solar eruptions.

4. Discussion

4.1. Why is the Special Update Required for the Vertical Velocity Located at $z = \Delta z$?

In the above section, we found that the case of CV-A0 does not capture the ground-truth data well; for instance, in the temporal evolution of the kinetic energy and the axis of the MFR. One other notable difference, when compared to the other cases from the data-driven simulation, is that the eruption occurs later in the CV-A0 case. The difference between these cases is the handling of the vertical velocity ($v_z$) at $k = 1$, i.e., $z = \Delta z$. Therefore, we should examine the behavior of the vertical velocity at $z = \Delta z$ or around the position, in both the energy buildup and energy release phases, respectively.

Figure 15 shows the temporal evolutions of the velocity field obtained from the ground-truth data and the data-driven MHD simulation (TW-A0 and TW-A2), respectively, at the surface at $z = 3\Delta z$. Although the horizontal velocity plotted by the arrows forms a twisting motion in each case, the vertical velocity ($v_z$) distributions are quite different. The most striking difference is that the negative vertical velocity appears in the late phase of the data-driven MHD simulations. So, hereafter we discuss the spatial and temporal evolutions of the vertical velocity.

Figure 16 shows the temporal evolutions of the vertical velocity along the center of the numerical box, i.e., the vertical dashed line as shown in Figure 14(a), which are obtained from ground-truth data and the data-driven simulations, respectively, in the energy buildup phase (TW-A0 to TW-A2). For the ground-truth data, the velocity temporally increases according to the twisting motion given on the photosphere. Although the vertical velocity in TW-A1 and TW-A2 increases as time goes on, it enhances the negative value with time in the region close to the solar surface, which is different from the ground-truth data. The typical result is the case of TW-A0 in which the velocity is negatively increased close to the solar surface with time. The small inset is an enlarged view in the height range of 0–0.01. We found that the velocity located at one grid above the bottom surface, i.e., $z = \Delta z$, is the fastest way to enhance the negative value. The time variation of the vertical velocity in the case of TW-A0 is determined by the equation of motion while in TW-A1 it is set to zero and in TW-A2 it is updated by...
linear interpolation in the vertical direction, respectively. Either the convective term $v \cdot \nabla v$ or Lorentz force $J \times B$, or both would have a negative influence on the vertical velocity. Despite each velocity profile beginning differently in each case of the energy buildup phase, magnetic energy is found to follow almost the same behavior in each case (Figure 8). Since the magnetic energy is dominant over the kinetic energy, the evolution of the magnetic field is almost unaffected by the velocity field; conversely, a small fluctuation of the magnetic field would greatly influence the evolution of the velocity. Therefore, the Lorentz force would be a major factor that inhibits the vertical velocity from leading to the correct solution. Note that, in the real situation, since no one can say that the magnetic energy dominates and the velocity field does not affect the magnetic evolution much on the photosphere, the situation is expected to become more complex.

Figure 17 plots the temporal evolution of the vertical velocity in the energy release phase, i.e., the eruption phase of the MFR in the same format as in Figure 16. The data-driven simulations, CA-V1 and CA-V2 capture the ground-truth data well while the results obtained from CV-A0 show the different behavior. The small inset follows the same format as in Figure 16(b), and the value at $z = \Delta z$ suddenly becomes the negative value. This behavior is the same as seen in the case of TW-A0 in the energy buildup phase. However, this case differs from TW-A0 in that the velocity recovers from its negative value during the late phase.

Figures 18(a) and (b) compare the temporal evolutions of the vertical velocity located at $z = \Delta z$ for the ground-truth data and
the results obtained from between the data-driven simulation in the energy buildup phase (TW-A0 and TW-A2) and energy release phase (CV-A0 and CV-A2). Note that the vertical velocity at $z = \Delta z$ for the cases of TW-A1 and CV-A1 is set to zero, so we exclude these plots. The TW-A2 and CV-A2 cases somewhat capture the ground-truth data, while cases TW-A0 and CV-A0 show large deviations from the ground truth by exhibiting a steep drop in the negative value from the initial values. CV-A0 however is found to increase the velocity back toward that seen in the ground-truth data. It is likely that when the magnetic field is converted into the dynamics phase from the static phase, the strong positive enhancement of the velocity is associated with the eruption, which returns to the ground-truth data even at $z = \Delta z$.

As can be seen from these results, we need a careful treatment of the vertical velocity at $z = \Delta z$ in the static phase, which is found to largely deviate from the ground-truth data without the treatments. Since the kinetic energy in the energy buildup phase is much smaller than the magnetic energy, the evolution of the magnetic field is unaffected by the velocity field even if the velocity strongly deviates from the ground-truth data as can be seen in Figure 8. On the other hand, it affects the initiation of the eruption, which causes delays because the downflow, which is an unlikely result in the ground-truth data, inhibits local reconnection at the photosphere and therefore the formation of the MFR. Thus, some treatments (simple linear interpolation was applicable in this study) would be required.

Why does the vertical velocity ($v_z$) located at $z = \Delta z$ show different behaviors in each case? From the above results, when the difference in the ground-truth data is striking, the magnetic field evolves in the quasi-statistical energy buildup phase toward the pre-eruption stage. Since the magnitude of the kinetic energy is much lower than the magnetic energy in the energy buildup phase, $J \times B$ is the major factor in lowering the velocity. Each component of the Lorentz force ($F$) is described...
where $B_z$ at $z = \Delta z$ is derived from the rotation of the electric fields located at $z = \Delta z$ while a derivative of $E_x$ and $E_z$ in the $z$ component is included to derive the $B_x$ and $B_y$, respectively, according to Equation (22). $B_z$ is determined by $E_{k=1}$ which is derived from the MHD electric field $(-v \times B)$ while $B_x$ and $B_y$ are derived from $E_{k=2}$, $E_{k=1}$, and $E^{(3)}$ too (see Equation (22)). $E_{k=2}$ and $E_{k=1}$ are different from $E^{(3)}$ because $E^{(3)}$ is derived from a non-MHD process. Therefore, the $z$ derivative is generally not allowed at $z = \Delta z$ and physical consistency of the obtained $B_x$ and $B_y$ is not guaranteed. $F_z$ includes more $B_x$ and $B_y$ than the other components $F_x$ and $F_y$, so the value of $v_z$ that is derived from $F_z$ will have a lesser accuracy.

4.2. What is the Difference between Type A and Type B?

We now discuss results obtained from the data-driven simulation for the case of TW-B, which implements the differential method Type B described in Equation (23). We first compare with the results obtained from the ground-truth data. Figures 19(a) and (b) plot the temporal evolutions of the
magnetic and kinetic energies in the energy buildup phase, i.e., the magnetic field is twisted due to the photospheric motion. Both evolutions obtained from the data-driven simulation show large deviations from the ground-truth data. Figures 19(c)–(e) show the distribution of $|J|/|B|$ plotted on the vertical cross section for the ground-truth data and the data-driven MHD simulations (TW-A2 and TW-B), respectively, from which the TW-A2 case reproduces the ground-truth data well, while the case of TW-B shows a totally different solution.

The difference between TW-A2 and TW-B is the application of differing numerical differential methods, especially, at $z = \Delta z$ at $k = 1$. The derivative forms in the $z$ direction for each and are as follows, with Type A written as

$$
\frac{\partial E}{\partial z}_{k=1} = \frac{1}{2} \left[ \frac{E_{k+2} - E_{k+1}}{\Delta z} + \left( \frac{E_{k+1} - E_0}{\Delta z} \right) \right],
$$

and Type B is written as

$$
\frac{\partial E}{\partial z}_{k=1} = \frac{(E_{k+1} + E_{k+1})}{2} - \frac{E_{k+1}}{\Delta z}. 
$$

Type A is an average value of the central differential method of electric fields obtained from the MHD process (the first term on the right-hand side of Equation (25)) and non-MHD process (the second term). Although each derivative has a physical meaning, there is a discrepancy between the MHD term and non-MHD term. The use of the average process in this case smoothly connects them, making a more robust calculation. Type B is a central differential method of the electric fields obtained from the MHD and non-MHD processes. In principle, this derivative is not allowed as discussed in the section above because the origin of those electric fields is different, i.e., there are not connected smoothly. Therefore, this discrepancy might negatively impact the simulation result.

On the other hand, Hayashi et al. (2018, 2019) implement Type B and the simulation works well. One reason for the difference in results from this study is the difference in the numerical schemes used. Hayashi et al. (2018, 2019) used a numerical scheme designed with the finite volume method, and numerical viscosity and resistivity might efficiently work on the negative impact and make the simulation stable. However, such viscosity and resistivity do not work efficiently in the central differential method as used in this study despite these
being included explicitly in the induction equation and the equation of motion. Therefore, numerical instability is not suppressed in this study and the solution is not reproduced correctly as can be seen in Figure 19 (e).

4.3. Dependency on Time Cadence

We test the effect on the accuracy of our results to the time cadence used in the data-driven MHD simulations used in this study. Figure 20(a) shows the comparison of the temporal evolution of the kinetic energies in the energy buildup phase that is obtained from the ground-truth data, the data-driven simulation (TW-A2), and a new data-driven simulation with 2.5 times the temporal resolution of TW-A2 (TW-A2D). These are plotted in black, red, and blue, respectively. Since the blue line almost overlaps with the red line, no significant differences due to the temporal evolution are apparent in the energy buildup phase. Thus, we can say that the solution is convergent with respect to temporal resolution.

Figure 20(b) shows each kinetic energy during the energy release phase, plotted in the same format as Figure 20(a). The red line, which corresponds to the data-driven simulation (CV-A2), diverges slightly from the blue line (CV-A2D), which has 2.5 times the temporal resolution of CV-A2. This difference is from the result in the energy buildup phase. However, the difference between CV-A2 and CV-A2D is only a few percent. From these results, our solution is almost convergent in temporal resolution. According to Leake et al. (2017), using a higher temporal cadence makes the results of a data-driven simulation closer to the ground-truth data. However, we do not see this in our results. One possible reason could be due to differences in the methods used to create the simulation. Leake et al. (2017) directly use data that are extracted from the ground-truth data whereas our method drives the coronal magnetic field by electric fields given near the solar surface that are derived from the bottom magnetic fields of the ground-truth data. It is therefore likely that these differing methods are the reason for the differing effect of adjusting the temporal resolution of the simulation.

Furthermore, it is important to consider a ratio of the sampling time ($dt_s$) to the dynamical time ($dt_d$) proposed by Leake et al. (2017) and the ratio should be less than 1. The dynamical time is defined by $dt_d = \Delta/V_p$, where $\Delta$ is the length of the grid interval and $V_p$ is the velocity of the plasma.
Figure 17. The temporal evolution of $v_z$ at the center of the numerical box associated with the converging motion given on the photosphere. The format is the same as Figure 16 where (a)–(d) correspond to the results obtained from the ground-truth data and the data-driven MHD simulations (CV-A0, CV-A1, and CV-A2). The black, purple, green, blue, and red lines represent the height profiles of the vertical velocities at $t = 22.50, 24.74, 26.99, 29.24$, and $31.49$, respectively. The small inset corresponds to the enlarged view in the height range of 0–0.01.

Figure 18. (a) Temporal evolution of $v_z$ at $k = 1$ is plotted for each case in the energy buildup stage. Each color, black, purple, and red, represents the results obtained from the ground-truth data and the data-driven simulations (TW-A0 and TW-A2), respectively. (b) Temporal evolution $v_z$ at $k = 1$ is plotted for each case in the energy release stage. Each color, black, purple, and red, represents the results obtained from the ground-truth data and the data-driven simulations (CV-A0 and CV-A2), respectively.
photosphere. In this study, since $\Delta = 1/320$ and maximum $V_p(\text{max}) = 0.01$, $dt_{\text{min}}$ is given as 0.3125. The sampling time $dt_x$ is given as 0.5625 (low-time resolution cases) and 0.225 (high-time resolution cases); therefore, the ratio corresponds to 1.8 and 0.72, respectively. The ratio in the low-time resolution cases (other than TW-A2D and CV-A2D) is over 1. However, in this work, the bottom-boundary parameters driving the system are from a model and thoroughly given in form of functions of time and position. For this reason, there is little uncertainty in the boundary data (sampling data) and there is not much difference in energies between the low-time resolution cases and high-time resolution cases.

4.4. Data-driven Simulation through the Energy Accumulation Process to the Release Process

We run the data-driven simulation throughout the entire process from the energy accumulation process to the release process. The result is shown in Figure 21(a) in which the

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure19.png}
\caption{(a) The temporal evolution of the magnetic energies obtained from the ground-truth data in black and the data-driven simulation (TW-B) in green. (b) The temporal evolution of the kinetic energies in each case, which is in the same format as (a). (c)–(e) The temporal evolution of $|J|/|B|$ obtained from the ground-truth data and the data-driven simulations (TW-A2 and TW-B), respectively.}
\end{figure}
The temporal evolution of the ground-truth data and the data-driven simulation is plotted in black and red, respectively. We found that the data-driven simulation cannot cause the eruption that is different from the results discussed above. One of the possible reasons is the pre-eruption magnetic field. Figures 21(b) and (c) show the pre-eruption magnetic fields of the ground-truth data and produced by the data-driven simulation, respectively. The size of the sigmoidal structure is different between them. Figure 10(b) shows the magnetic flux of the pre-eruption magnetic fields that depends on the twist value before the

Figure 20. (a) The temporal evolutions of the kinetic energy in the twisting phase for the ground-truth data in black, the data-driven simulations (TW-A2 and TW-A2D) shown in red and blue, respectively. TW-A2D has 2.5 times the temporal resolution of TW-A2. (b) The temporal evolutions of the kinetic energy in the eruption phase for each case. The format is the same as that shown in (a).

Figure 21. (a) The temporal evolution of the kinetic energy from $t = 0$ to the last time of the simulation for the ground-truth data shown in black and the data-driven simulation shown in red, respectively. This data-driven simulation was conducted in the same manner in TW-A2 and CV-A2. (b) Three-dimensional magnetic field lines in the pre-eruption phase at $t = 23.26$ for the ground-truth data. (c) Three-dimensional magnetic field lines in the pre-eruption phase for the data-driven simulation.
relaxation process. From this result, the maximum twist value of the ground-truth data is in the range of 0.6–0.7. On the other hand, the highly twisted field lines buildup in the data-driven simulation are concentrated approximately up to $T_u = 0.6$. This may be very small to cause the eruption. For instance, Inoue et al. (2018) suggested that causing the solar flares associated with the eruptions requires the twisted field lines to have twist values from 0.5–1.0. Therefore, the ground-truth data does not have much of a twist and it would be difficult for the magnetic field produced by the data-driven simulation to cause the eruption. This result indicates that a little difference in twist may affect the eruption. Furthermore, this is a reminder that the 3D solution is not uniquely determined even using the same boundary condition.

Although the problem may be caused by the ground-truth data that has weakly twisted field lines before the eruption in this study, we should avoid the problem of weakening of the twist of the magnetic field lines produced in the data-driven simulation when using the photospheric magnetic field. The NLFFF would be useful as the initial condition of the data-driven simulation before the eruption because in some cases, the NLFFF before the eruption reproduces more highly twisted field lines than those shown in Figure 21(c) (Inoue 2016). However, further discussion is needed on the issue of whether to use the magnetic field produced in the data-driven simulation or the NLFFF as the initial condition. This will be addressed in future work.

5. Summary

The three-dimensional coronal magnetic field plays an essential role in producing solar phenomena such as solar flares, solar jets, and CMEs. However, the 3D coronal magnetic field is not fully understood as only observations of the photospheric magnetic field can currently be made. Data-driven simulations (Toriumi et al. 2020) are a powerful tool to detect not only the 3D magnetic structure but also chart its evolution. However, an important question about their reliability still remains. Therefore, in this study, we developed a new data-driven MHD code based on that of Hayashi et al. (2018) which is designed with a zero-beta approximation and a central differential method, and confirmed its performance. In order to check the performance of the developed data-driven MHD simulation code, we carried out a benchmark test using ground-truth data made from a time-evolving 3D magnetic field produced in the MHD simulation (Amari et al. 2000 and Amari et al. 2003). This covers the typical subsequent evolution from the energy buildup stage to the eruption stage, as seen in a flare-productive active region (Toriumi & Wang 2019). Leake et al. (2017) and Toriumi et al. (2020) tested the accuracy of their data-driven simulations using the ground-truth data that was made from flux emergence simulations. In this study, our ground-truth data includes the eruption process of the MFR. Consequently, we confirmed that our data-driven simulation can reproduce the eruption processes of the magnetic field well.

We found that if there are no treatments for the vertical velocity ($v_z$) at $z = \Delta z$, the temporal evolution deviates strongly from the ground-truth data. This is inferred from the discrepancy due to a derivative in the $z$ derivation of the electric field, which includes an MHD electric field ($-\nabla \times B$) and a non-MHD electric field ($E^{\parallel}$). In this study, we found that simple linear interpolation or rather $v_z$ set to zero better suppresses the deviation from the original MHD solution. We found that these deviations are larger when the magnetic field evolves in a quasi-static energy buildup process toward the eruption. After the eruption, we found that the vertical velocity returns toward the values seen in the ground-truth data. There are no remarkable differences between the magnetic field evolutions, which are obtained from the data-driven simulations, in the energy buildup phase even though the vertical velocity at $z = \Delta z$ in each case is not the same way. However, from Figure 11(b), the initiation of the MFR eruption is later than other cases obtained from the data-driven simulation if the vertical velocity does not behave correctly. This result suggests that the downflow, which is unlikely in the ground-truth data, inhibits reconnection at the photosphere and hence the formation of the MFR (see Figure 13(a)). Therefore, we require the treatment of the vertical velocity at $z = \Delta z$.

There are several choices for the discretization methods applied at $z = \Delta z$, which is the practical boundary condition of the method proposed in Hayashi et al. (2018). The critical issue is the derivative in the $z$ direction between the electric fields obtained from the MHD and non-MHD processes. This is because, in general, this derivative is not allowed and that contradiction causes the numerical instability. We suggest that the stability strongly depends on the numerical scheme used. In addition, we here note that the vertical size of the simulation grid ($\Delta z$) can be arbitrarily determined in the simulation setup; therefore, changing $\Delta z$ can be a remedy to reduce the differences in the simulation results of Types A and B. We will investigate this point further in the future.

In this study, we found that our data-driven simulation could capture an overview of the evolution of the 3D coronal magnetic field, which mimics the evolution as seen in the flare-productive active region. As a next step, we will apply the observed photospheric magnetic field to our data-driven simulation code. In general, there is no guarantee that the success of this simulation will result in the success of this future simulation. For example, in this study, we assigned enough spatial resolution to the PIL, where the converging motion drives the eruption. Although it is important to correctly capture the photospheric motion near the PIL for the eruption, some observational data do not provide that. Recently, large ground-based telescopes such as the GST and the Daniel K. Inouye Solar Telescope (Rast et al. 2021) have provided high-resolution data, which will be helpful in resolving this issue. To use these high-resolution data, we will need further optimization and to implement further technical upgrades in our code. Additionally, in this study, we did not take into account $v_z$ at the bottom surface in the ground-truth data as in Leake et al. (2017) and Toriumi et al. (2020), so we have not estimated how much this velocity affects the data-driven simulation. However, we believe that our data-driven simulation has the potential to reproduce the evolution of the magnetic field from the energy buildup stage to its eruption stage even using the observed magnetic field based on its performance in this study.

We are grateful to a referee for constructive comments and Dr. Magnus Woods for reading this paper. This work was supported by the National Science Foundation under grant Nos. AGS-1954737, AGS-2145253 and AST-2204384 and the National Aeronautics and Space Administration under grant No. 80NSSC21K1671. This work was also supported by the computational joint research program of the Institute for Space-
Appendix A

Derivation of the Poisson Equations

A.1. Derivation of the Second Poisson Equation for $\Phi^{(2)}$

First, we multiply $\nabla_h$ by Equation (12), then

$$\nabla_h \cdot (\nabla \times E^{(2)}) = \frac{\partial}{\partial z} \left( \frac{\partial E^{(2)}_x}{\partial x} + \frac{\partial E^{(2)}_y}{\partial y} \right) = 0,$$

should be satisfied at the boundary. If we assume

$$E^{(2)} = -z \Phi^{(2)}, \quad (A1)$$

the above relationship is satisfied. We substitute Equation (A1) into Equation (12), then we can obtain

$$- \nabla \times E^{(2)} = - \nabla \times (-z \Phi^{(2)}) = \left( \frac{\partial \Phi^{(2)}}{\partial y}, -\frac{\partial \Phi^{(2)}}{\partial x}, 0 \right).$$

Therefore, we describe the equation for $B_{h,df}$ as follows:

$$\frac{\partial B_{h,df}}{\partial t} = \left( \frac{\partial \Phi^{(2)}}{\partial y}, -\frac{\partial \Phi^{(2)}}{\partial x}, 0 \right). \quad (A2)$$

We multiply $\nabla \times$ and its $z$ component corresponds to the second Poisson equation (Equation (15)),

$$\nabla^2 \Phi^{(2)} = -z \cdot \left( \nabla \times \frac{\partial B_{h,df}}{\partial t} \right)$$

$$= -\left\{ \frac{\partial}{\partial x} \left( \frac{\partial B_{h,df}}{\partial t} \right) - \frac{\partial}{\partial y} \left( \frac{\partial B_{h,df}}{\partial t} \right) \right\}$$

$$= -\left\{ \frac{\partial}{\partial x} \left( \frac{\partial B_{h,df}}{\partial t} \right) - \frac{\partial}{\partial y} \left( \frac{\partial B_{h,df}}{\partial t} \right) \right\}. \quad (A3)$$

Since the right-hand term can be derived from the observed $B_x$ and $B_y$, we can find $\Phi^{(2)}$ and eventually obtain $E^{(2)}$ through Equation (A1). Note that we used

$$\frac{\partial}{\partial x} \left( \frac{\partial B_{h,df}}{\partial t} \right) - \frac{\partial}{\partial y} \left( \frac{\partial B_{h,df}}{\partial t} \right) = 0.$$

A.2. Derivation of the Third Poisson Equation for $\Phi^{(3)}$

The solenoidal condition, $\nabla \cdot B = 0$ can be rewritten as follows:

$$\nabla_h \cdot B_h + \frac{\partial B_z}{\partial z} = 0,$$

and we obtain the following relationship by conducting the time derivative of the above equation,

$$\frac{\partial}{\partial z} \left( \frac{\partial B_z}{\partial t} \right) = -\nabla_h \cdot \left( \frac{\partial B_{h,df}}{\partial t} \right)$$

$$= -\nabla_h \cdot \left( \frac{\partial B_{h,df}}{\partial t} \right)$$

$$= -\frac{\partial}{\partial z} \left( \nabla \times E^{(3)} \right) \cdot \mathbf{z}. \quad (A3)$$

Equation (A3) can be described by taking into account Equation (17), as follows:

$$\frac{\partial B^{(3)}_h}{\partial t} \bigg|_{z=+\frac{1}{2}\Delta z} - \frac{\partial B^{(3)}_h}{\partial t} \bigg|_{z=0} = -\frac{1}{2} \Delta z \nabla_h \cdot \left( \frac{\partial B_{h,df}}{\partial t} \right). \quad (A4)$$

When we make the same assumption as when we obtained $E^{(1)}$, $E^{(3)} = z \times \nabla_h \Phi^{(3)} = \left( -\frac{\partial \Phi^{(3)}}{\partial y}, \frac{\partial \Phi^{(3)}}{\partial x}, 0 \right)$, (A5) the equation

$$\nabla \times E^{(3)} = \left( \frac{\partial^2 \Phi^{(3)}}{\partial y^2}, \frac{\partial^2 \Phi^{(3)}}{\partial x^2}, -\frac{\partial B^{(3)}_h}{\partial t} \right). \quad (A6)$$

is obtained. Eventually, we can obtain the third Poisson equation,

$$\nabla^2 \Phi^{(3)} = \frac{1}{2} \Delta z \nabla_h \cdot \left( \frac{\partial B_{h,df}}{\partial t} \right).$$

Appendix B

Performances of the Poisson Equations

We show the performances of the Poisson Equations (14)–(16) by evaluating the following values:

$$L_1 = \int \left| \nabla^2 \Phi^{(1)} \right| dV,$$

$$L_2 = \int \left| \nabla^2 \Phi^{(2)} \right| dV,$$

$$L_3 = \int \left| \nabla^2 \Phi^{(3)} - \frac{1}{2} \Delta z \nabla_h \cdot \left( \frac{\partial B_{h,df}}{\partial t} \right) \right| dV, \quad (B1)$$

and

$$\text{ER}_1 = \left| \nabla^2 \Phi^{(1)} \right|_{\text{max}} - \left| \frac{\partial B}{\partial t} \right|_{\text{max}},$$

$$\text{ER}_2 = \left| \nabla^2 \Phi^{(2)} \right|_{\text{max}} - \left| \frac{z \cdot \left( \nabla \times \frac{\partial B_{h,df}}{\partial t} \right)}{\text{max}} \right|,$$

$$\text{ER}_3 = \left| \nabla^2 \Phi^{(3)} \right|_{\text{max}} - \left| \frac{1}{2} \Delta z \nabla_h \cdot \left( \frac{\partial B_{h,df}}{\partial t} \right) \right|_{\text{max}}. \quad (B2)$$
When the solutions of each Poisson equation are correctly obtained, these values completely drop to zero. The Poisson equation is solved numerically by a simple Gauss–Seidel method based on the second-order finite difference method and the initial $\Phi$ is given as zero. Figures A1(a) and (b) show the iteration profile of $L_n$ and $ER_n$, respectively, where $n = 1, 2, 3$. We find that each value dramatically decreases during each iteration and eventually saturates at a very low value compared to the initial value. Therefore, the solutions could be obtained with good accuracy.

**Appendix C**

**Reproducibility of the Photospheric Magnetic Field**

We check how much reproducibility of the photospheric magnetic field from the electric field through Equation (20) has. Following Hayashi et al. (2018), the bottom magnetic field should be reproduced perfectly through the data-driven simulation but it depends on the accuracy of the Poisson solver. Following them, we also make scatter plots for the photospheric magnetic field of the ground-truth data versus the magnetic field reproduced via the data-driven simulation using the electric fields. Figure A2 shows scatter plots for $B_x$ and $B_y$ in the twisting phase from $t = 0 \sim t = 11.25$. Note that the $B_z$ component is not plotted because this does not change during this time period. These plots are almost along a function, $y = x$, i.e., the magnetic field is reproduced well from the data-driven simulation. Figure A3 shows these scatter plots in the eruption phase, which are in the same format as in Figure A2 except $B_z$ is plotted. Although there are little dispersions compared to previous cases (however, the correlation coefficient is over 0.99), the scatter plots are mostly along the line of $y = x$. So, throughout the simulation, the boundary magnetic field is reproduced well from the data-driven simulation based on the given electric fields.
Figure A2. Scatter plot of the bottom magnetic fields by the ground-truth data vs. those reproduced through the data-driven simulation. The upper and lower panels correspond to the distributions of $B_x$ and $B_y$, respectively, at each time under the twisting motion. Note that the scatter plots on $B_z$ are excluded because it does not change during the evolution.
Figure A3. Scatter plots of the bottom magnetic fields from the ground-truth data vs. that reproduced through the data-driven simulation during the eruption stage, which are presented in the same format as in Figure A2. The upper, middle, and lower panels are the results of the $B_x$, $B_y$, and $B_z$ components, respectively. Although these look a little scattered compared to Figure A2, the correlation coefficients are over 0.99.

ORCID iDs
Satoshi Inoue https://orcid.org/0000-0001-5121-5122
Keiji Hayashi https://orcid.org/0000-0001-9046-6688
Takahiro Miyoshi https://orcid.org/0000-0002-4675-4460

References
Amari, T., Luciani, J. F., Aly, J. J., et al. 2003, ApJ, 585, 1073
Amari, T., Luciani, J. F., Mikic, Z., et al. 2000, ApJL, 529, L49
Berger, M. A., & Prior, C. 2006, JPhA, 39, 8321
Bobra, M. G., Sun, X., Hoeksema, J. T., et al. 2014, SoPh, 289, 3549
Cheung, M. C. M., & DeRosa, M. L. 2012, ApJ, 757, 147
Clyne, J., Mininni, P., Norton, A., et al. 2007, NJPh, 9, 301
Clyne, J., & Rast, M. 2005, Proc. SPIE, 5669, 284

Dedner, A., Kemm, F., Kröner, D., et al. 2002, JCoPh, 175, 645
Fisher, G. H., Kazachenko, M. D., Welsch, B. T., et al. 2020, ApJS, 248, 2
Fisher, G. H., Welsch, B. T., Abbott, W. P., et al. 2010, ApJ, 715, 242
Goode, P. R., & Cao, W. 2012, Proc. SPIE, 8444, 844403
Guo, Y., Xia, C., Keppens, R., et al. 2019, ApJL, 870, L28
Hayashi, K., Feng, X., Xiong, M., et al. 2018, ApJ, 855, 11
Hayashi, K., Feng, X., Xiong, M., et al. 2019, ApJL, 871, L28
Inoue, S. 2016, PEPS, 3, 19
Inoue, S., Hayashi, K., Magara, T., et al. 2014, ApJ, 788, 182
Inoue, S., Kusano, K., Büchner, J., et al. 2018, NaCo, 9, 174
Inoue, S., Kusano, K., Magara, T., et al. 2011, ApJ, 738, 161
Inoue, S., Magara, T., Watari, S., et al. 2012, ApJ, 747, 65
Jiang, C., Bian, X., Sun, T., et al. 2021, FrP, 9, 224
Jiang, C., Feng, X., Guo, Y., et al. 2022, Innov, 3, 100236
Jiang, C., Feng, X., & Hu, Q. 2018, ApJ, 866, 96
Jiang, C., Wu, S. T., Feng, X., et al. 2014, ApJ, 780, 55
