


FULL PAPER

Open Access



An alternate representation of the geomagnetic core field obtained using machine learning

Lukács Kuslits^{1*} , András Horváth², Viktor Wesztergom¹, Ciaran Beggan³, Tibor Rubóczki¹, Ernő Prácser¹, Lili Czirok⁴, István Bozsó¹ and István Lempenger¹

Abstract

Machine learning (ML) as a tool is rapidly emerging in various branches of contemporary geophysical research. To date, however, rarely has it been applied specifically for the study of Earth's internal magnetic field and the geodynamo. Prevailing methods currently used in inferring the characteristic properties and the probable time evolution of the geodynamo are mostly based on reduced representations of magnetohydrodynamics (MHD). This study introduces a new inference method, referred to as Current Loop-based UNet Model Segmentation Inference (CLUMSI). Its long-term goal focuses on uncovering concentrations of electric current densities inside the core as the direct sources of the magnetic field itself, rather than computing the fluid motion using MHD. CLUMSI relies on simplified models in which equivalent current loops represent electric current systems emerging in turbulent geodynamo simulations. Various configurations of such loop models are utilized to produce synthetic magnetic field and secular variation (SV) maps computed at the core–mantle boundary (CMB). The resulting maps are then presented as training samples to an image-processing neural network designed specifically for solving image segmentation problems. This network essentially learns to infer the parameters and configuration of the loops in each model based on the corresponding CMB maps. In addition, with the help of the Domain Adversarial Training of Neural Networks (DANN) method during training, historical geomagnetic field data could also be considered alongside the synthetic samples. This implementation can increase the likelihood that a network trained primarily on synthetic data will appropriately handle real inputs. Our results focus mainly on the method's feasibility when applied to synthetic data and the quality of these inferences. A single evaluation of the trained network can recover the overall distribution of loop parameters with reasonable accuracy. To better represent conditions in the outer core, the study also proposes a computationally feasible process to account for magnetic diffusion and the corresponding induced currents in the loop models. However, the quality of the reconstruction of magnetic field properties is compromised by occasional poor inferences, and an inability to recover realistic SV.

Keywords Geomagnetic field, Geodynamo, Current loops, Physics informed neural networks, Domain adversarial training

*Correspondence:

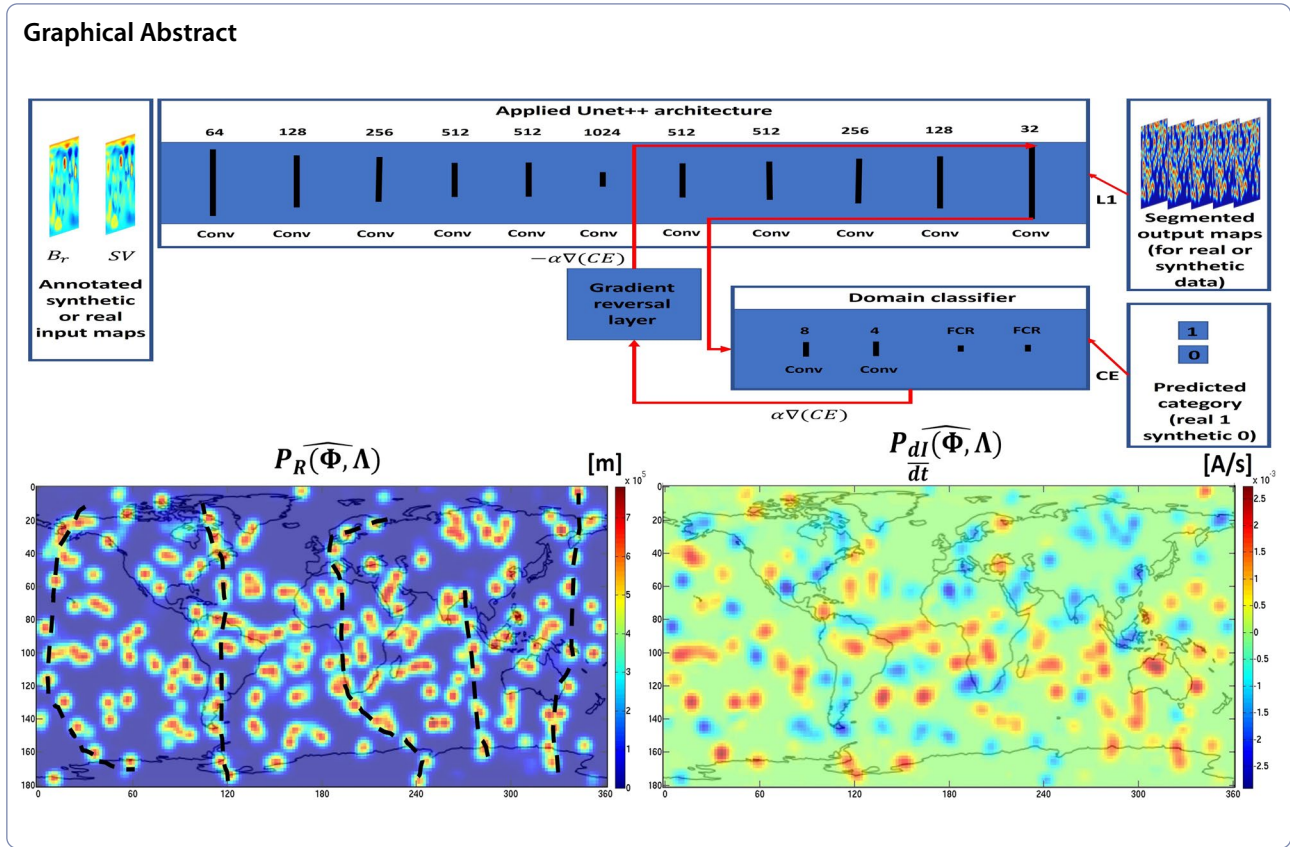
Lukács Kuslits

kuslits.lukacs@epss.hun-ren.hu

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.



Introduction

There are currently four main research directions aiming at exploring how the geodynamo operates. Core surface flow inversions, inverse geodynamo modeling, laboratory experiments and direct numerical simulations (Glatzmaier and Olson 2005; Christensen 2011). The first two of these are the most prevalent tools at present used for inferring the actual state of the geodynamo (Huder et al. 2019). To do this, they make use of the dimensionless form of the induction Eq. (1) describing the MHD interaction between the fluid flows and magnetic fields:

$$\partial_t \mathbf{B} = \nabla \times (\mathbf{u} \times \mathbf{B}) + \frac{1}{Rm} \nabla^2 \mathbf{B}. \tag{1}$$

In Eq. (1), \mathbf{u} contains the flow velocities, \mathbf{B} is the magnetic induction. The dimensionless magnetic Reynolds number $Rm = UL/\eta$ is used to represent the ratio of forces between the flow-driven advection of the magnetic field and the viscous dissipation of the field in the absence of flow (also known as magnetic diffusion (Holme et al. 2015)). In the magnetic Reynolds number, L is a characteristic length scale, U is the typical speed of fluid flows and $\eta = 1/(\mu_0\sigma)$ is the magnetic diffusivity, μ_0 and σ are the magnetic permeability of free space and the electrical conductivity, respectively.

Core surface flow inversions aim to reconstruct horizontal flow patterns and velocities at the core–mantle boundary (CMB) with the observation of the radial component of the magnetic field using Eq. (1). This approach suffers from non-uniqueness issues arising from the ill-posed nature of the solution and its uncertainty along lines of zero radial induction (Whaler 1986). Interestingly, even such uncertain solutions can lead to more accurate forecasts of SV over three to five years than estimates produced by linear extrapolations of the time dependency of the Gauss coefficients (Whaler and Began 2015).

Inverse geodynamo modeling makes use of the quasi-geostrophic approximation of the already simplified Boussinesq equations of convection attached to Eq. (1) to simulate a dynamic interaction of induced magnetic fields and fluid flow in a conductive core (Gillet et al. 2011). These types of dynamical models have the beneficial property of a coupling between flow velocities at the boundary layer (CMB) and those in the bulk of the core. This allows for a Bayesian inference of the distribution of the internal physical parameters of the system from surface magnetic field and flow data which can be fed back into the simulations for predictive modeling using a technique called variational data assimilation

(Talagrand 1997). The disadvantage of this technique lies in the instabilities of ensemble modeling and the complex statistical properties (covariances) of simulated parameters being relied on during computations (Sanchez et al. 2018).

Laboratory experiments are perhaps the least commonly investigated of the above-mentioned methods of inquiry. Nevertheless they have provided a handful of results relevant with respect to the geodynamo (e.g., Müller et al. 2008; Miralles et al. 2013; Monchaux et al. 2010; Su et al. 2020). These experiments can also provide valuable basis for testing the robustness and quality of computational reconstruction methods, providing ground truth data about what is expected from complex MHD processes such as the geodynamo.

Direct numerical simulations on the other hand aim at developing computational models which can reproduce the turbulent dynamic regime expected for the core dynamo as authentically as possible. Though this goal is severely hampered in particular by the low viscosity values estimated for the outer core, which seem impossible to approach even exploiting state of the art computational power (Sheyko et al. 2016; Dong et al. 2021), significant leaps towards it have been made. Recent examples of this include the models of (Aubert et al. 2017) and (Aubert 2023). Interestingly, in the rare case where in such studies electric current densities were also reported, results display concentrated current systems emerging with increasing vigorousness. This concentration manifests itself in coil-like current systems surrounding sheet-like plumes of axisymmetric flow (Miyagoshi et al. 2008). The general picture one can obtain about the distribution of electric currents in such simulations is that a significant part of field generation is carried out by these localized current systems appearing with relatively uniform geometries. It is to be noted that the sheet plumes characterizing flow velocities in a rapidly rotating turbulent regime associated with the above-mentioned current systems appear in other published work on simulations results, such as Schaeffer et al. (2017) or Aubert (2019).

The geodynamo simulations raise the intriguing prospect of producing simplified models based on idealizing the geometries of these concentrated currents. To do so might be beneficial when one would like to gain information on regions inside the geodynamo which can act as sources of the geomagnetic field without any direct need for the computational complexity necessary to conduct a simulation approximating real Earth-like circumstances. This line of thought can only be valid of course provided that the overall structure of the turbulent MHD regime and the associated current density distribution in the simulations indeed resembles that of the geodynamo, and if some practical way of inference can be established

using such idealized phenomenological models. In the current study, one possible way of bypassing this second obstacle with the help of ML is presented in detail.

Until now, few studies investigating the internal geomagnetic field have exploited ML methods. For example, the work of (Gwirtz et al. 2022) concerns the predictability of pole reversals using ML only on historical geomagnetic dipole moment (GMD) data. The authors concluded that the task could not be well addressed by the methods they studied due to the small number of relevant data and their low frequency domain resolution. It is worth mentioning Loftin et al. (2019) who analyzed the applicability of ML in the data preparation for geomagnetic models.

In an introductory study Kuslits et al. (2020) gave a more detailed review of the difficulties faced by current research directions not using ML, and proposed a very similar alternative method to the one presented here. Conceptually, the forward model idealized the geodynamo process as a set of many localized individual current loops (or loops for short), and the inversion method heavily relied on ML, in particular, deep learning. However, that study demonstrated the concept in one relatively simple example using only synthetic data with no conductive medium and no time variation introduced in the forward model. It was presented as being an equivalent loop model with an emphasis on being conscious of the much more complicated current systems potentially existing in various spatial scales within a conductive core in the actual geodynamo.

It is to be noted, that reconstructing a complex current density distribution in a large inaccessible volume of space like Earth's outer core (in which the bulk of the dynamo action occurs) has its specific challenges.

There is a long history of other previous works searching for current loop representations of the geomagnetic field with a similar goal in mind, such as Alldredge (1987), Rong et al. (2021) and references therein and Peddie (1979). All these attempted to fit either one or more magnetic dipoles or current loops to some representation of the Earth's magnetic field or a dataset that contains measured magnetic field values. Computationally speaking, they generally inverted for loop parameters using some form of least-squares inversion. Although in the case of a single current loop the determination of loop parameters through least-squares inversion yields sound results, such parameters rarely provide insight into the inner workings of the geodynamo or the current systems in the Earth's outer core. It is possible to attempt to model the Earth's magnetic field by carrying out least-squares inversion for the parameters of multiple loops. In this case, the sensitivity of least-squares inversion to the initial parameters can be remedied in multiple ways.

In Allredge (1987) the author first inverted the parameters of two loops in axial position. Then based on the residual magnetic field, judiciously set the initial parameters for 5 additional current loops and carried out the inversion for the newly added loop parameters. These steps were iterated until sufficiently low RMS error was reached. In Peddie (1979), the author ran 20 sets of inversions with randomized initial conditions to account for the sensitivity to initial parameters.

To summarize and expand on the shortcomings of these attempts, the following observations can be made:

- 1) They applied little to no objectively grounded constraints on the possible characteristic dimensions of the loops or their abundance in the models.
- 2) They offered no objective methodological possibility to handle the discrepancies between idealized representations, and the potentially much more complex magnetic fields and current systems existing within the geodynamo.
- 3) The linearized least-squares inversions they mostly applied were sensitive to the initial parametrization of the loops, in particular, to the initial placement of them (Gubbins and Roberts 1987). The solutions offered by different authors relied on either a human expert to determine sound initial parameters or the randomization of the initial parameters. These solutions made the regular and “large-scale” application of such inversion methods untenable (e.g., inverting for different IGRF representations).
- 4) They did not take into account “secondary” effects when attempting to reconstruct the time-variation of the field as well as its actual state, such as the dampening of the magnetic field (magnetic diffusion) generated by changes of the loop currents in time.

As it was also learned from previous attempts using other types of equivalent source approaches for forward modeling, such as Mayhew and Estes (1983) and Ladynin (2014), these problems are necessary to deal with when one aims at developing a physically relevant reconstruction of the geodynamo using highly simplified sources.

The present work addresses these shortcomings via upgrading the approach proposed in the introductory study of Kuslits et al. (2020). This further developed approach is referred from hereafter as the CLUMSI methodology. Shortcomings were addressed specifically the following ways. Issue (1), by deriving the range of the potential number of loops and loop parameter values. Issues (2) and (3) by training an updated image-processing deep neural network using the DANN method, which then gives an initial estimation for all the loop parameters corresponding to a set of field and SV values over

the course of a single evaluation. Issue (4) by establishing a practical approximation based on a systematic series of simple numerical models to consider the effect of electromagnetic induction on time-varying magnetic fields of loop models in a highly conductive medium such as Earth’s core.

Figure 1 presents the flowchart showing an overview of the CLUMSI estimation scheme, components of which are explained mainly in the following two sections.

Section “[Defining an idealized model and constraining the model parameters](#)” gives the detailed description of the forward model (components [1–7]). Section “[Inversion framework](#)” gives a high-level summary of the inversion framework concentrating on the applied deep learning image segmentation algorithm along with introducing the loss functions and quality parameters used during training and evaluation (components [8–12], [16]).

Section “[Testing the methodology](#)” summarizes inversion results using noise-free synthetic data and demonstrates a result on input data coming from a real geomagnetic model (components [13–15], [17]). Discussing the pros and cons of CLUMSI and the implications of these results on some ideas concerning further development is featured in Sect. “[Discussion](#)”. Conclusions are given in Sect. “[Conclusions](#)”. The appendices detail further aspects of the inversion framework, especially concerning the representativeness of training samples. One concrete example demonstrating why the type of problem defined in this study is more difficult to handle using a previously applied least-squares inversion technique is presented in Appendix G.

Defining an idealized model and constraining the model parameters

To formulate the forward problem in a way which satisfies the desired phenomenological concept drafted in Sect. “[Introduction](#)”, a simplified current system model is needed, in which individual sources can meet the qualitative criteria described below.

The structure and features should roughly correspond to those of the current systems emerging in the Miyagoshi et al. (2008) and Miyagoshi et al. (2011) simulations. They should also have a finite spatial extent, while forming a closed circuit and must be well parametrized. After iterating on possible geometric configurations, filamentary circular currents (loops) were chosen as the simplest possible current systems which can fulfill these requirements.

As mentioned earlier, for the analysis presented here to be more relevant for the actual geodynamo process, the stationary current loop models used in previous studies needed to be updated. To do that, the solution applied

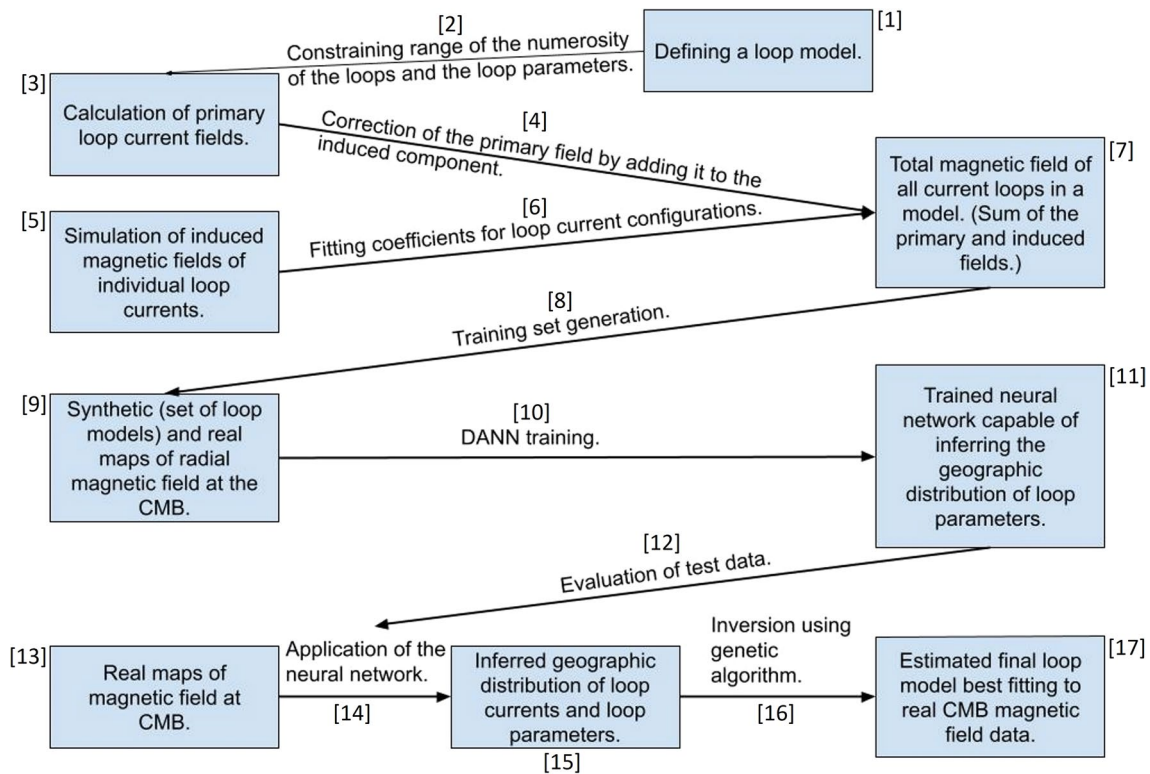


Fig. 1 Flowchart of the entire algorithmic scheme of the estimation process presented in this study

Table 1 Parameters of a single current loop

Notation [unit]	Description of source parameter
Φ_i [°]	Geographic co-latitude of the center of the loop
Δ_i [°]	Geographic longitude of the center of the loop
r_i [m]	Distance from the center of the Earth to the center of the loop
R_i [m]	Loop radius
$I_i(t)$ [A]	Intensity of the current carried by the loop (time dependent)
θ_i [°]	Declination of the loop axis
λ_i [°]	Azimuth of the loop axis
$(dI/dt)_i$ [A/s]	Rate of change of the current intensity in time

in Kuslits et al. (2020) to compute the field of stationary loops and is described in detail in Sect. “Primary field of stationary current loops” was corrected by an estimate of the additional (induced) field produced by a time-variation in the loop currents (see Sect. “Approximating the total field assuming linear time variation in the source currents”).

Altogether each loop in this updated model can be described by 8 source parameters featured in Table 1.

For the solutions to be eligible to produce a training set representative of historical geomagnetic field data, additional constraints based on first-order considerations

about the geomagnetic field were imposed when generating the model samples (see Appendix B).

Primary field of stationary current loops

The radial component of the primary magnetic field of an individual current loop i is obtained in the Cartesian coordinate system according to the following analytical solution deduced by Simpson et al. (2001).

Provided that the current loop has an axial position in its local coordinate system (see Fig. 2a) and the field is computed at an observation point located at a distance

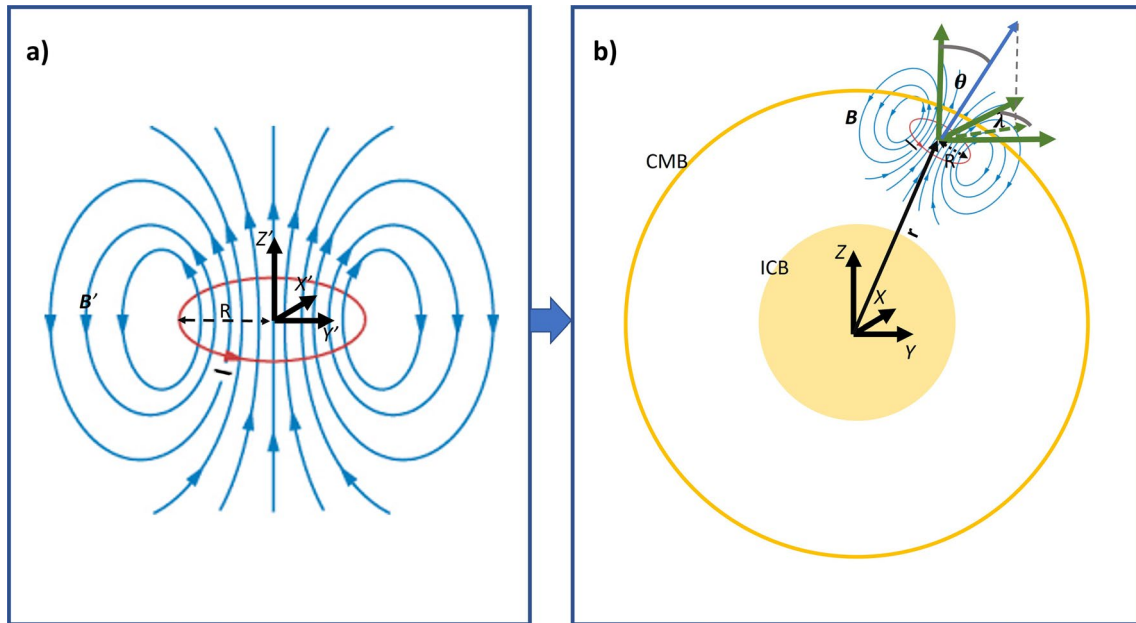


Fig. 2 Schematics of transforming a given current loop from a source-centered (local, **a**) Cartesian coordinate system to an Earth-centered (global, **b**) Cartesian coordinate system in the models. θ and λ are the angles of attack in transformations (7, 8) and define the attitude of a given loop. (Source of the figure used for the magnetic field lines: (Ling et al. 2016))

$\mathbf{r} = (dx_i, dy_i, dz_i) = (x - x_{0i}, y - y_{0i}, z - z_{0i})$ from the center of the loop:

$$B_{x_i}(\mathbf{r}) = \frac{CI_i x_i z_i}{2\alpha_i^2 \beta_i \rho_i^2} \left[(R_i^2 + r_i^2) E(k_i^2) - \alpha_i^2 K(k_i^2) \right], \tag{2}$$

$$B_{y_i}(\mathbf{r}) = \frac{dy_i}{dx_i} B_{x_i}, \tag{3}$$

$$B_{z_i}(\mathbf{r}) = \frac{CI_i}{2\alpha_i^2 \beta_i} \left[(R_i^2 - r_i^2) E(k_i^2) + \alpha_i^2 K(k_i^2) \right]. \tag{4}$$

In the above expressions, $\rho_i^2 = dx_i^2 + dy_i^2$, $r_i^2 = dx_i^2 + dy_i^2 + dz_i^2$, $CI_i = \frac{\mu_0 I_i}{4\pi}$, $\alpha_i^2 = R_i^2 + r_i^2 - 2R_i \rho_i$, $\beta_i^2 = R_i^2 + r_i^2 + 2R_i \rho_i$ and $k_i^2 = 1 - \frac{\alpha_i^2}{\beta_i^2}$. $E(k_i^2)$ and $K(k_i^2)$ are elliptic integrals of the first and second kind, respectively.

The final result for the radial magnetic field was calculated for the CMB as reference surface using a Mercator projection and it was obtained with the help of the following series of spatial rotations.

Positional rotation of the source vector field by a declination angle θ_i :

$$\mathbf{R}_{\theta_i} = \begin{bmatrix} \cos(\theta_i) & 0 & \sin(\theta_i) \\ 0 & 1 & 0 \\ -\sin(\theta_i) & 0 & \cos(\theta_i) \end{bmatrix}. \tag{5}$$

Positional rotation of the source field by an azimuth angle λ_i :

$$\mathbf{R}_{\lambda_i} = \begin{bmatrix} \cos(\lambda_i) & -\sin(\lambda_i) & 0 \\ \sin(\lambda_i) & \cos(\lambda_i) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{6}$$

Using matrices (5) and (6), one can determine the spatial coordinates of the magnetic induction vector field of each i unique source:

$$[x_i, y_i, z_i] = (\mathbf{R}_{\lambda_i} \mathbf{R}_{\theta_i})^{-1} \begin{bmatrix} x'_i \\ y'_i \\ z'_i \end{bmatrix}. \tag{7}$$

The same transformation must be performed for the spatial attitudes of induction vector fields to obtain magnetic fields corresponding the spatial orientation of each source at each point on the CMB surface:

$$[B_{x_i}, B_{y_i}, B_{z_i}] = (\mathbf{R}_{\lambda_i} \mathbf{R}_{\theta_i})^{-1} \begin{bmatrix} B'_{x_i} \\ B'_{y_i} \\ B'_{z_i} \end{bmatrix}. \tag{8}$$

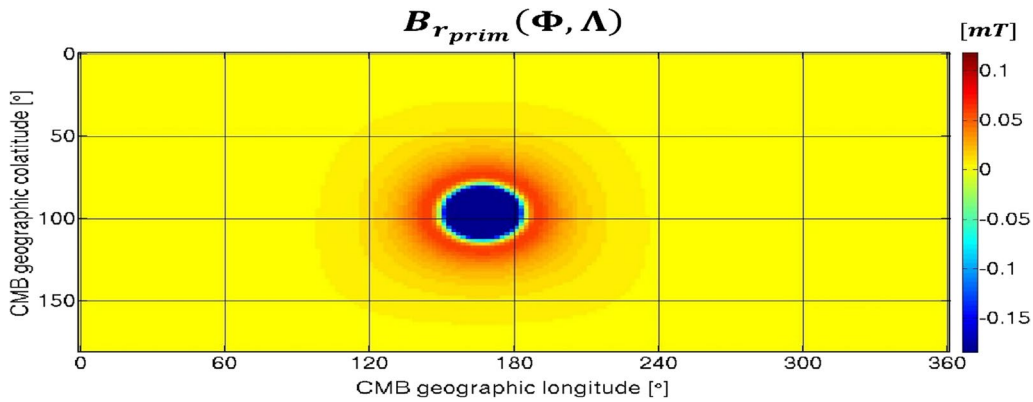


Fig. 3 Primary radial magnetic field of a single, radially aligned source ($\phi = 95^\circ$, $\Lambda = 165^\circ$, $r = 3e6$, $R = 8.4e5$, $l = 8.4e8$, $\theta = 95^\circ$, $\lambda = 165^\circ$) computed for a grid with a resolution of 2° (90 by 180 gridpoints) on the reference surface of the models representing the CMB

The radial component of the resulting magnetic field in spherical coordinate system was then obtained from the Cartesian components at each (CMB) surface point:

$$B_{r_{\text{prim}}}^i(r, \phi, \Lambda) = B_{x_i} \sin(\phi) \cos(\Lambda) + B_{y_i} \sin(\phi) \sin(\Lambda) + B_{z_i} \cos(\phi). \quad (9)$$

The solution for the primary radial field in the forward problem uses the summation of radial field component data at the CMB for all the N sources introduced in a particular model instance. As a result, Mercator maps (such as the one in Fig. 3) of the primary magnetic field of the entire model on the CMB were obtained for each model. Due to limitations in computational capacity, an angular resolution of 2° in longitude and latitude was set when producing the resulting maps:

$$B_{r_{\text{prim}}}(r = r_{\text{CMB}}, \phi, \Lambda) = \sum_{i=1}^N B_{r_{\text{prim}}}^i(r = r_{\text{CMB}}, \phi, \Lambda). \quad (10)$$

Approximating the total field assuming linear time variation in the source currents

Due to the high estimated values for core conductivity (Ohta et al. 2016; Dongxiao et al. 2019), the effect of magnetic diffusion cannot be neglected when approximating the total magnetic fields of the source models.

The conductive material of the core and the strongly time-dependent MHD processes necessitated the introduction of a model encompassing time-varying sources. Such a model also considers the magnetic field of currents induced in a highly conductive medium and is thus much more relevant for the geodynamo than applying solely the solution discussed in Sect. “Primary field of stationary current loops” which assumes only steady-state loops.

However, to construct this model in a consistent way while still being able to produce sets of solutions quickly

enough to be able to apply ML in the inverse problem mounted a significant challenge.

To generate sufficient training data, a systematic series of simulations of the induced fields around individual loops applying different source parameters and current variations was built up. Electromagnetic induction was modeled by a spherical domain representing Earth’s core in the simulations. As the exact value of electrical conductivity in the core is relatively poorly constrained, the authors chose a conservative value of $\sigma = 5 * 10^5$ [S/m] (this value was selected based mainly on the experimental work of Ohta et al. (2016)). In each simulation the \mathbf{B}_{prim} primary field and the \mathbf{B}_{tot} total field of the source were computed separately. The induction effect was given by the difference between these fields ($\mathbf{B}_{\text{tot}} - \mathbf{B}_{\text{prim}}$). These simulations used a finite element numerical framework (see Fig. 4) utilizing the COMSOL Multiphysics 5.3a software package (Multiphysics 1998).

To be able to later produce enough training data for the ML-based parameter inference, the following simplifying assumptions need to be made:

- 1) The location of the current loops is fixed, and two loops cannot fall directly under each other in the radial direction. The former of these two conditions meant that one did not have to account for induction coming from the motion of the sources in the conducting medium. The latter constraint was intended to prevent sources that are geographically very close to each other which leads to problems of equivalence (see Sect. “Results using noise-free synthetic data”).
- 2) The position of the sources is set, so that the axes are radial (that is for all i , $\Phi_i = \theta_i$; $\Lambda_i = \lambda_i$).
- 3) Only the current intensity carried by the sources varies, and it varies linearly in time.

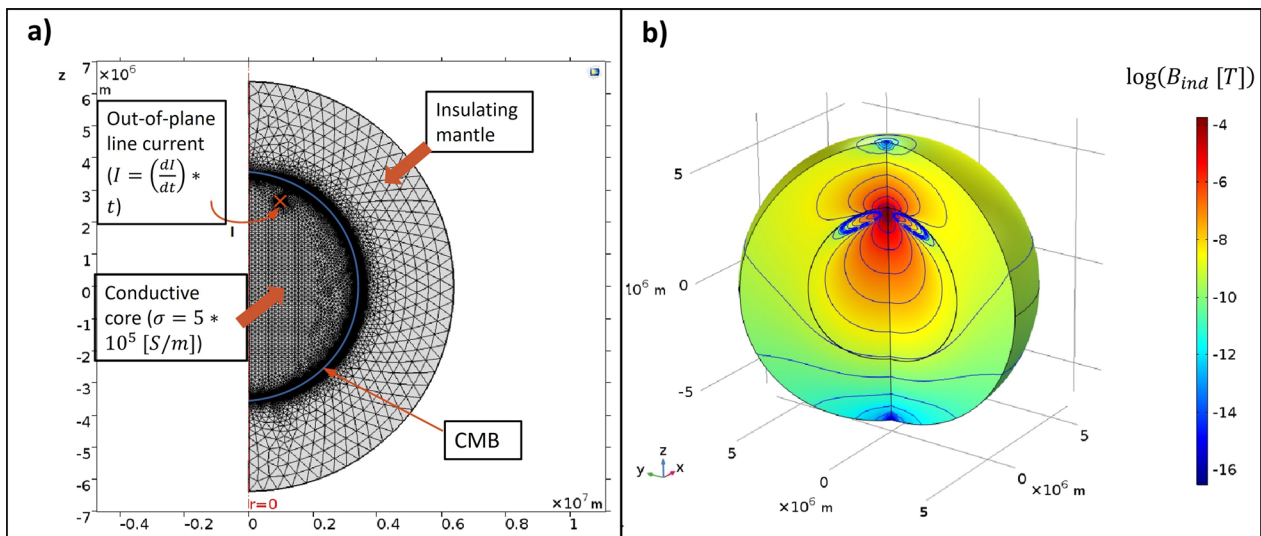


Fig. 4 Computational setup of a single simulation member in the series showing the finite element mesh (a) and the resulting induced field around the loop in the rotationally symmetric spherical domain (b)

4) The inner and outer core of the Earth were modeled as one uniformly conducting sphere. A full description of this approach is given by Metman et al. (2019).

Conditions (1)–(2) allowed us to perform a series of simulations in an axisymmetric 2D domain. This procedure is also known as 2.5D simulation, since it is possible to perform 3D modeling by exploiting the symmetry property of a 2D geometric layout (Jacobs et al. 2007).

Condition (4) simplified the setup and computational complexity of the simulations.

The application of condition (3) also had a beneficial consequence in terms of computational complexity. This meant that the solution of a complex frequency domain induction problem calculating the field of a source emitting harmonic signals (Weaver 1994) could be avoided. As a result, induced fields independent of time and the magnitude of the primary source could be obtained. As the transients decayed, the magnitude of induced fields depended linearly on the rate of change in the source currents. The current intensity in the models can be described by the form shown below:

$$I = I_{prim} + I_{ind} = C(r)t + I_{ind}(r, t). \tag{11}$$

In the above equation, the primary current I_{prim} comes from a current loop and changes linearly in time at a given rate $C(r)$. This constraint on the current results in a linearly changing magnetic field in time around the source, decreasing with distance r from the source depending on the electrical conductivity of the surrounding medium. As the current loop is embedded within a conductive medium, representing the Earth’s core, the

time-variation of the magnetic field induces a complex current density field around it. Induced currents $I_{ind}(r, t)$ have a direction at each point in any given time which decreases the change in the magnetic flux responsible for their creation. This results in a screening effect by the conductive core depending on the shape and size of the conductive domain between the source and each point of observation. Following the decay of transients, static-induced fields proportional only to $C(r)$ are formed.

This equilibrium state thus sets in after some delay in simulation time with an increasing spherical distance from the source on the CMB. In our series of simulations, following a delay time of $t_d = 10^6 [yr]$, these static-induced fields have built up in all our models on the CMB.

Using these results meant that a polynomial interpolation of the above-described induction effect (Eqs. (19–23)) using data points computed in the systematic numerical simulations was satisfactory with respect to generating synthetic training samples within computational (and time) limitations.

The numerical simulations used a low-frequency approximation of Maxwell’s equations in time domain. Separate simulations computing the primary and the total magnetic fields were run in parallel, assuming an insulating and a conductive core, respectively. They were implemented using a spherical axisymmetric geometry introducing current loops with systematically varied parameters described in Table 1. This meant the numerical solution of equations:

$$B = \nabla \times A, \tag{12}$$

$$\nabla \times \mathbf{B} = \mu \mathbf{J}, \tag{13}$$

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t}, \tag{14}$$

$$\mathbf{J} = \sigma \mathbf{E}. \tag{15}$$

For clarity, in Eqs. (12) and (14), \mathbf{A} denotes the magnetic vector potential and \mathbf{E} the electric field. To set the dimensions of the simulation domain to be characteristic for those of Earth, the following radii were defined:

$$r_{\text{CMB}} = 3.48 * 10^6 [\text{m}], r_E = 6.38 * 10^6 [\text{m}], \tag{16}$$

r_{CMB} denoting the radius of the CMB and r_E that of the Earth.

Current loops are defined as line-currents perpendicular to the plane in which we calculate induced field values in this axisymmetric setting. Since the magnitude of static induced fields depended linearly on the rate of change in the source currents, interpolation can be carried out using only two values of $(dI/dt)_i$ independently from the current intensities (which were chosen to be zero at the beginning of the simulation time for the sake of simplicity).

Thus, three of the eight parameters listed in Table 1 were enough to obtain a basis set of full spatial solutions using individual current loops parametrized as described in Table 2. These were chosen from the range of potential loop parameters defined in Appendix B.

The boundary condition prescribed the vanishing of tangential magnetic fields, which is not ideal as outside the core poloidal magnetic fields can be expected (Metman et al. 2019). Although for this type of problem, applying

$$\mathbf{n} \times \mathbf{H} = 0 \tag{17}$$

at the surface of the Earth rather than at the CMB (where \mathbf{n} is the normal vector with respect to Earth's surface) was acceptable, as solutions in this case agreed to a high degree of accuracy regardless of the specific type of boundary condition used.

The magnetic effect of currents induced in a conductive core was determined only with respect to the radial magnetic component of induction vector fields at the CMB. As the components of the magnetic induction vector were defined in a cylindrical coordinate system, spherical radial components were computed as follows:

Table 2 Parameters of simulation members (loops), forming a total set containing 130 base results for the interpolation using all parameter combinations $r_i, R_i, \left(\frac{dI}{dt}\right)_i \in [|r| \times |R| \times |dI/dt|]$

Sets of varied loop parameters

$$|r|[m] = \{3.4e5, 3.2e5, 3e5, 2.8e5, 2.6e5\}$$

$$|R|[m] = \{2e5, 2.5e5, 3e5, 3.5e5, 4e5, 4.5e5, 5e5, 5.5e5, 6e5, 6.5e5, 7e5, 7.5e5, 8e5\}$$

$$|dI/dt|\left[\frac{A}{s}\right] = \{1e-3, 1e0\}$$

$$\begin{aligned} B_{r_{\text{ind}}}^i(r, \phi', t) &= B_{\rho_{\text{tot}}}^i(r, \phi', t) \sin(\phi') \\ &+ B_{z_{\text{tot}}}^i(r, \phi', t) \cos(\phi) \\ &- B_{\rho_{\text{prim}}}^i(r, \phi', t) \sin(\phi') \\ &- B_{z_{\text{prim}}}^i(r, \phi', t) \cos(\phi') |r \\ &= r_{\text{CMB}}; t = t_d. \end{aligned} \tag{18}$$

The relationship between these radial components and the corresponding loop parameters was then approximated fitting exponential polynomial basis functions using the weighted Ridge regression computation in the scikit-learn library (Pedregosa et al. 2011):

$$B_{r_{\text{ind}}}^i(r = r_{\text{CMB}}, \phi') \approx C_{\text{poly}}(R_i, r_i, \phi') * (dI/dt)_i, \tag{19}$$

where

$$C_{\text{poly}}(R_i, r_i, \phi') = \sum_{j=1}^{\binom{n+3}{n}-1} \beta_j e^{R l_{1j} r^{l_{2j}} \phi'^{l_{3j}}}. \tag{20}$$

In Eq. (20) $l_{1j} + l_{2j} + l_{3j} = n$ and $l_{1j}, l_{2j}, l_{3j} \geq 0$ and the β_j coefficients were determined from

$$\beta = \left(A^T W A + \chi I \right)^{-1} A^T W (\mathbf{B}_{\text{ind}} - \widehat{\mathbf{B}}_{\text{ind}}), \tag{21}$$

where I is the identity matrix, χ is the regularization parameter (see (Hoerl and Kennard 1970)), which was set to $\chi = 10^{-7}$. Matrix A contains the exponential of the loop- and angular distance parameters for the total number of $m = |R \times r \times \left(\frac{dI}{dt}\right) \times \phi'| = 65000$ simulated data points used for the approximation:

$$A_{m,j} = e^{R_m^{l_{1j}} r_m^{l_{2j}} \phi_m^{l_{3j}}}, \tag{22}$$

and W weights have the elements,

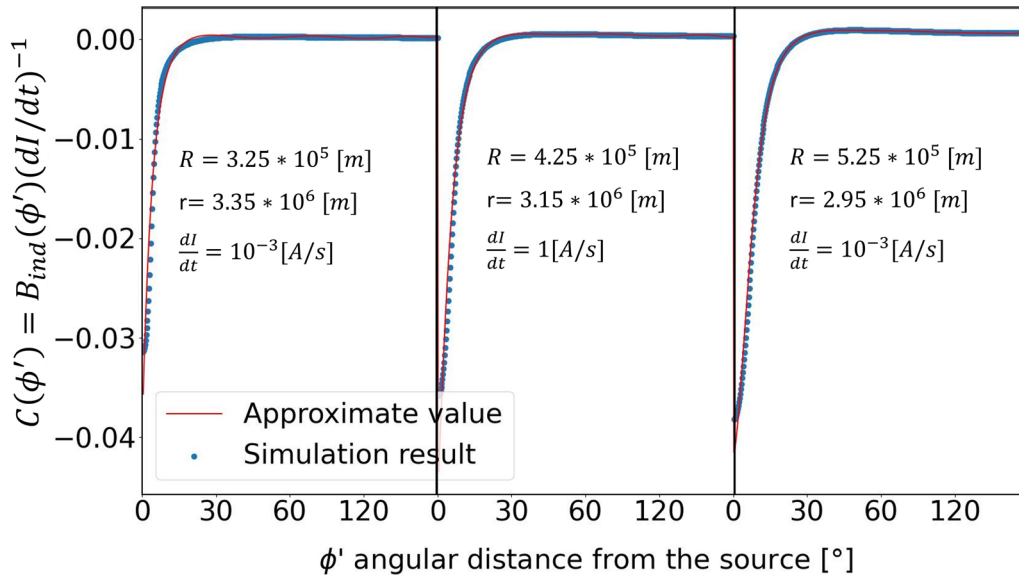


Fig. 5 Comparison of polynomials fitted for the induction effect on the CMB radial fields around three different example current loops and the actual results from test simulations. Source parameters are noted in the middle of each corresponding graph

$$W_{m,j} = \frac{1}{\sqrt{|B_{m,r_{ind}}^j|}}. \quad (23)$$

Using $n = 11$ and the weights as defined in (23), these exponential functions resulted in the best approximation of the original $B_{r_{ind}}^i$ with a normalized root mean square error (NRMS, see Sect. “Loss functions and quality metrics”) of $NRMS = 1.4 * 10^{-5}$ for the interpolated values (see the fitted polynomials in Fig. 5). The rationale behind applying approximation (19–23) is simply computational feasibility. Running individual simulations directly to compute the composite induced field of a single loop model containing a hundred loops, such as the ones shown in Fig. 6e and f would take approximately 15 min to compute. Using the polynomial formula (19–20) reduces the time needed for such a computation to less than 30 s.

Based on Eqs. (9–10) and (19–20), the final result can be obtained using both the primary and the induced magnetic fields for each current loop in an individual model sample. The spatial distribution of the magnetic fields can be computed relatively fast following the transformation of the axisymmetric solutions from the local coordinate systems of the loops to the models’ global spherical coordinate system:

$$\begin{aligned} B_{r_{tot}}(r = r_{CMB}, \phi, \Lambda) &= \sum_{i=1}^N \left(B_{r_{prim}}^i(r = r_{CMB}, \phi, \Lambda) \right. \\ &\quad \left. + B_{r_{ind}}^i(r = r_{CMB}, \phi_i' = a \cos(\mathbf{r} \cdot \mathbf{r}_{0i} / (r \cdot r_{0i}))) \right) \\ &= \sum_{i=1}^N B_{r_{prim}}^i(r = r_{CMB}, \phi, \Lambda) \\ &\quad + B_{r_{ind}}^i(r = r_{CMB}, \phi, \Lambda). \end{aligned} \quad (24)$$

In Eq. (24) \mathbf{r}_{0i} points to the center of the i th loop from that of the Earth, and \mathbf{r} points to a location on the CMB ($r = r_{CMB}, \phi, \Lambda$). These solutions yielded modified versions of the Mercator maps coming from solution Eq. (10) corrected for the screening effect of the induction.

The above-detailed concept was efficient in generating a large amount of training data which were representative in terms of the overall distribution and magnitude of the radial geomagnetic field at the CMB. However, it also had an important drawback when it comes to reproducing Earth-like SV values. As an assumption of linear time variation in the source currents results in a constant $B_{r_{ind}}^i$, any time variation in these synthetic magnetic fields is observed in their primary radial components $B_{r_{prim}}^i$. Source positions were fixed in the models so one can estimate what is the highest possible rate of change in the current carried by the largest loop, based on maximum observable values of the actual SV (see in Appendix B). Such a high rate-of-change would in turn produce an

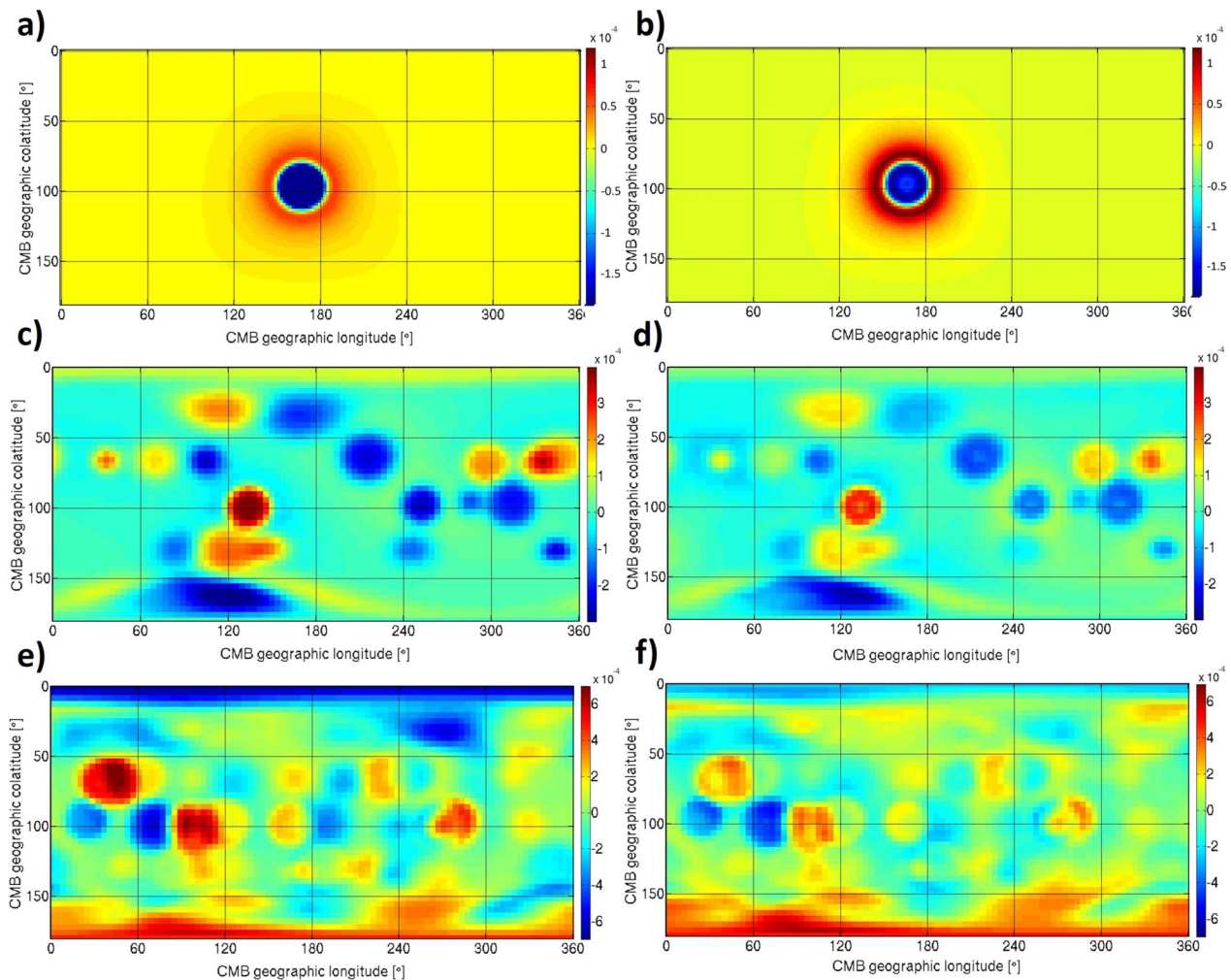


Fig. 6 Primary (a), (c), (e) and total (b), (d), (f) radial magnetic fields obtained using the approximate solution correcting for inductive screening at the CMB (in [7] units). Maps are shown for models containing, respectively, 1, 25 and 100 radially aligned current loops placed arbitrarily inside the domain representing an insulating (left column) and a conductive (right column) core

induced field so strong it would be an order of magnitude larger than the largest values of the actual radial CMB field recorded in available historical geomagnetic data (see Sect. “Refining the ML-based inversion” and Appendix C). This essentially means that the source model described here was unable to simultaneously reproduce Earth-like field magnitudes and SV. We chose to optimize the models for the former objective, so that training samples could be accepted as it is described in Appendix C.

Figure 6 demonstrates the general range of full model solutions used in the synthetic component of the training set and shows how induction screens the primary magnetic fields in the models. One can observe that even though the loop axes in these models are all radially aligned, more complex fields begin to form when increasing the number of loops in the models with the effect of individual loops becoming progressively more difficult to

separate for the human eye. Moreover, even a relatively subtle screening around the sources could also have a rather complex effect on the field maps, when comparing the original maps on the left to the corrected ones on the right-hand side of the figure.

Inversion framework

To come up with potentially physically meaningful results in case of real geomagnetic data, besides the regularization of the forward model presented in Sect. “Defining an idealized model and constraining the model parameters” and Appendix B, modifications on the previous inversion framework were needed. This affected the machine learning implementation, i.e., the algorithm used for training and the generation of the training dataset. The framework used by CLUMSI is similar to the one used in the

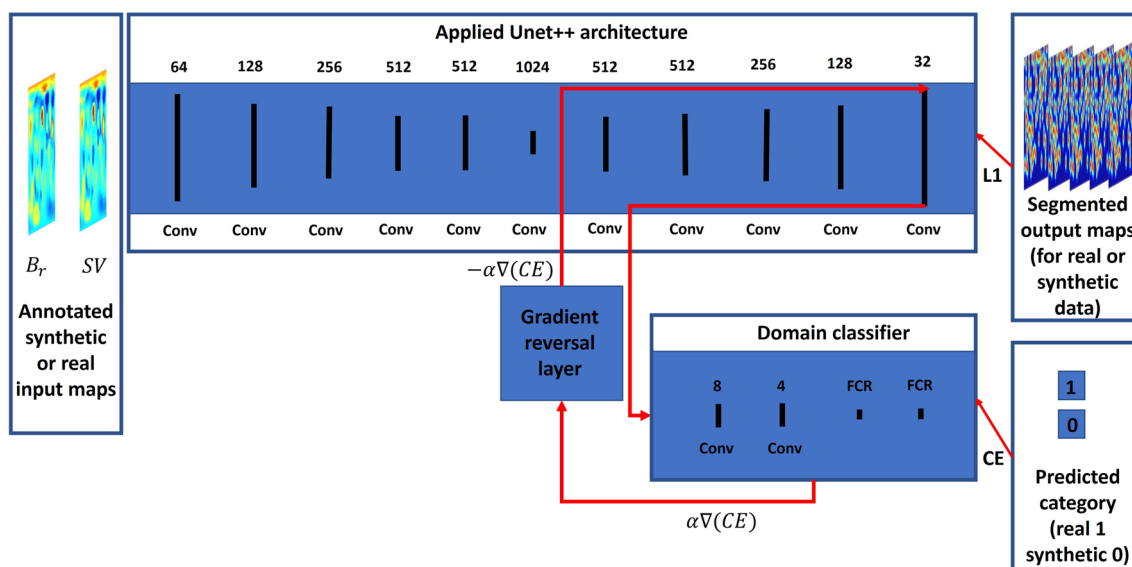


Fig. 7 Schematic diagram of the DANN training process of the refined neural network architecture working with complex source model data and real geomagnetic model data. Here, Conv stands for convolutional and FCR for fully connected layers. L1 denotes the loss function (25) and CE denotes the values of the cross-entropy (26) loss function. $\alpha \nabla(CE)$ denotes the weighted gradients of the cross-entropy function

introductory study in the sense that a neural network is trained to infer the distributions of source parameters (such as their geographic positions) represented as rectangular maps, from magnetic field maps used as input data. The most significant novel work undertaken is using information from measured geomagnetic data which was incorporated in the training process in such a way that allows the network to handle real data more correctly (via the DANN method).

Section “Refining the ML-based inversion” gives an overview of the applied inversion framework focusing on the deep learning methodology as the backbone of CLUMSI. Section “Loss functions and quality metrics” summarizes the quality measures applied for analyzing the method’s performance. A more detailed review of training and test data are presented in Appendix C. Appendix E describes the Genetic Algorithm (GA) as the last stage of CLUMSI obtaining a final estimation of the loop parameters and the corresponding magnetic fields based on the neural net inference.

Refining the ML-based inversion

To robustly obtain a fully parametrized reconstructed source model, a two-step inversion framework was implemented, supplemented by significant modifications when compared to the original algorithm in the introductory study.

That work suggested training an image segmentation neural network with a UNet architecture, only for detecting the geographic distribution of the sources using

the maps of the training set. This was followed by a GA solution which obtained exact value estimates for all the parameters of every loop from the suggested locations.

The current work replaces the original UNet implementation with a UNet++ architecture (Zhou et al. 2018) utilizing more hidden layers and output channels to give inferred images of all the source parameters in the models using their respective distribution maps as target values defined by Eqs. (60–63) in Appendix C. A new training method, referred to in Sect. “Introduction” as DANN (Ganin et al. 2016), was also implemented, which allowed for an efficient incorporation of maps coming from real geomagnetic data into the training set. Geomagnetic field model data used during training included radial field and SV maps from the historic data set of the GUFM-1 model spanning a time range from 1600 to 1975 (Jackson et al. 2000, see the data in Appendix C).

The refined training process and neural net architecture are drafted in a similar fashion to the flowchart of Ganin et al. (2016) in Fig. 7.

During the DANN training, the image segmentation task is augmented by a decision problem about whether the data are coming from the synthetic or the real component of the training set. This is achieved using a separate neural network called a domain classifier. The classifier is trained on the internal representations calculated in a hidden layer of the UNet++ network. That part of the training process is structured in such a way that the poorer the classifier performs, the better training results are achieved in terms of relying on features in

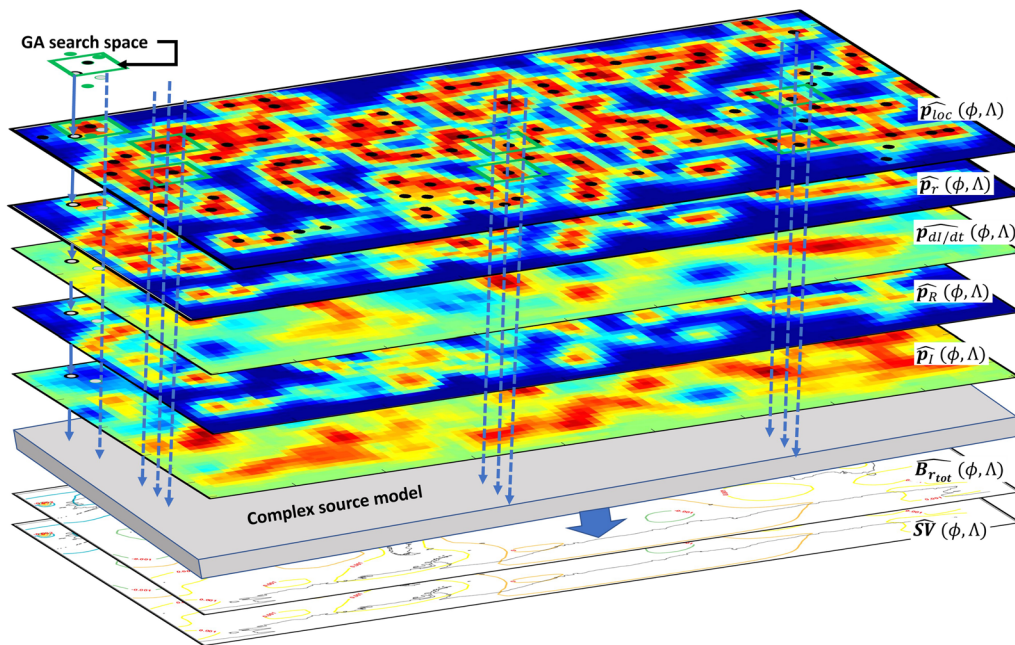


Fig. 8 Schematic diagram illustrating the final search process of the GA by ‘scanning’ the distribution functions estimated by the neural network. The black filled circles represent the selected initial source positions and the green squares the search space defined around them. The bottom two maps are the estimation results of the complex source model for Br and SV resulting from a given ‘scan’

the synthetic model which are similar to those of original geomagnetic data.

Finally, the GA solution was computed using the parameter distribution maps resulting from the UNet++ phase as inputs. The GA solution comprised a real coded, multi-population GA for deriving an optimally fitting parametrized current loop model from the neural net output maps. Figure 8 describes the procedure in which the GA searched only in a subspace of potential parameter values outlined around peaks in the parameter maps produced by the UNet++ evaluation. To pinpoint an estimated source location, these peaks were selected as local maxima on the map in the same fashion as described in (Kuslits et al. 2020). The GA search was then performed within a 2-gridpoint radius around them (in accordance with the constraint defined in Sect. “Approximating the total field assuming linear time variation in the source currents”). The quantities on the right-hand side of the figure, representing the parameter distribution maps, are introduced in detail in Appendix C. A more detailed description of the GA itself is given in Appendix E.

This seemingly complex solution described above not only improved the results when using synthetic data, compared to the first experimentation showed in the

introductory study, but also produced more relevant and well-defined solutions in case of real geomagnetic field data.

Figure 9 demonstrates the difference made by the DANN training and the modified UNet++ configuration on the resulting geographic distribution maps of source positions. It shows that when the previous and the new DANN trained networks are evaluated on the same input maps, the outputs significantly differ in definitiveness. This means not just a sharper outline of reconstructed probable source areas but a more precise solution quantifiable in case of synthetic input data as shown in Fig. 10 in Sect. “Results using noise-free synthetic data”.

Loss functions and quality metrics

When training UNet++ for the image segmentation, the L1 loss function was applied to all the channels providing the expected outputs to every source parameter distribution:

$$L1 = \sum_{j=1}^{n_{\text{data}}} \left| P_{k_j}(\phi, \Lambda) - \widehat{P_{k_j}}(\phi, \Lambda) \right|, \tag{25}$$

where n_{data} denotes the number of data points in a given sample, $P_{k_j}(\phi, \Lambda)$ and $\widehat{P_{k_j}}(\phi, \Lambda)$ are given values of each

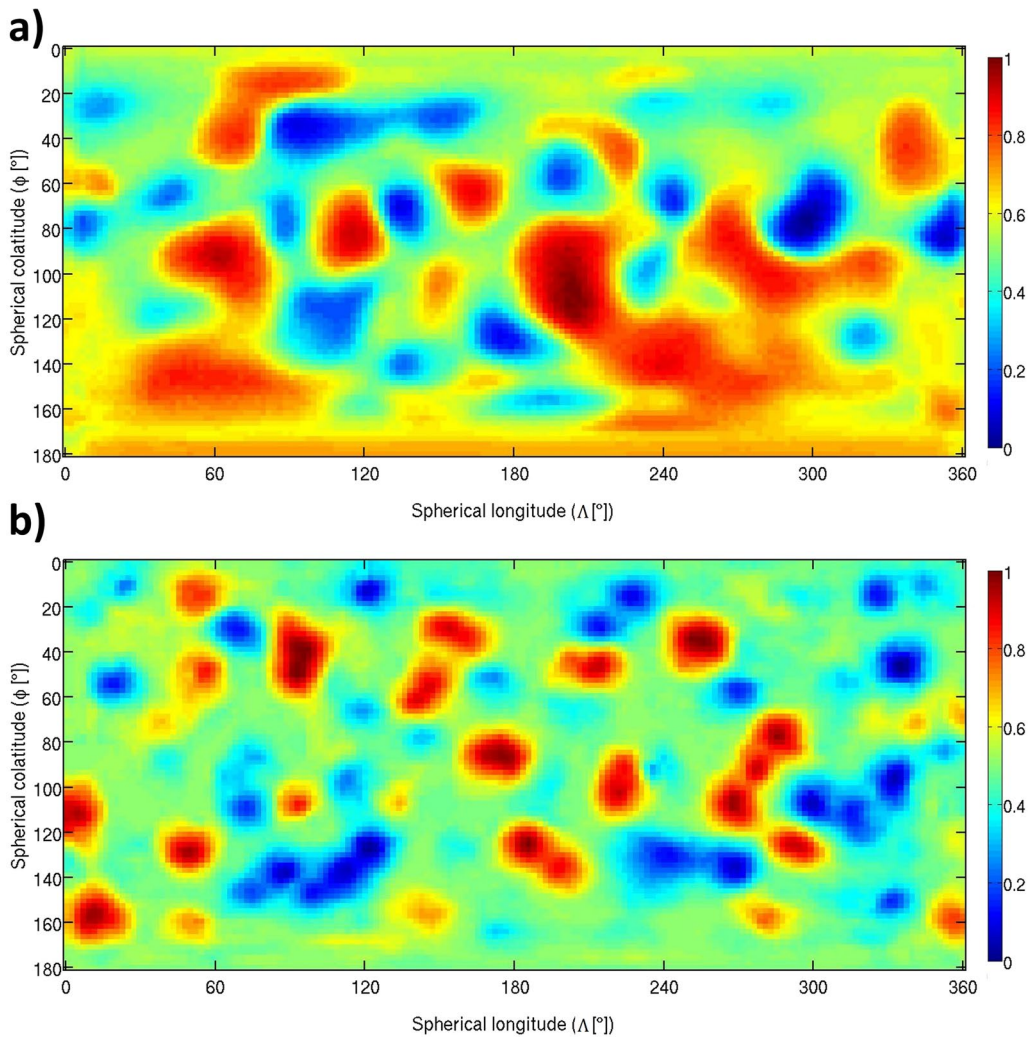


Fig. 9 Conventional (a) and DANN trained (b) neural network evaluation on real geomagnetic data (normalized inferred dl/dt source parameter distribution projected onto the CMB, calculated from GUFM-1 model data for year 1600)

data point j on maps of target and estimated source parameters, respectively, defined in Appendix C.

For training the domain classifier to determine the origin of each model sample l , we used the cross-entropy loss as defined below:

$$CE = -\sum_{l=1}^{n_{\text{samples}}} V_l \log(p_l), \tag{26}$$

where n_{samples} denotes the number of samples within a given part of the set of maps used for training, V_l is the

class indicator and p_l is the estimated probability for each sample l .

As it was established in Sect. “Approximating the total field assuming linear time variation in the source currents” we chose to optimize for fitting to the field values rather than to SV values. This meant quantifying the misfit between the actual radial magnetic field and SV maps and those computed using the full parameter estimation derived from the GA search as follows:

$$\overline{\text{NRMS}} = \frac{\sqrt{(n_{\text{data}})^{-1} \sum_{j=1}^{n_{\text{data}}} (B_{r_j} - \widehat{B}_{r_j})^2}}{\max(B_r) - \min(B_r)} + 0.1 \frac{\sqrt{(n_{\text{data}})^{-1} \sum_{j=1}^{n_{\text{data}}} (SV_j - \widehat{SV}_j)^2}}{\max(\text{SV}) - \min(\text{SV})} \tag{27}$$

Similarly, the mean absolute error (MAE) was also given to show the misfits' absolute value:

$$\text{MAE} = n_{\text{data}}^{-1} \sum_{j=1}^{n_{\text{data}}} |B_{r_j} - \widehat{B}_{r_j}|. \quad (28)$$

Relative parameter error (RPE) was used to describe the quality of the full parameter reconstruction resulting from the GA search by comparing the true loop parameters with those of the physically closest estimated loops:

$$\text{RPE} = \frac{\left(\sum_{i,k} \frac{|\widehat{\text{Par}}_i^k(|\widehat{r}_i - r_{\text{gt}_i}| = \min) - \text{Par}_{\text{gt}_i}^k|}{|\text{Par}_{\text{gt}_i}^k|} \right)}{8L}, \quad (29)$$

where $\text{Par}_{\text{gt}_i}^k$ is the ground truth value of a single parameter k of a loop i in a vector \mathbf{Par} containing each current loop parameters (see Table 1), $\widehat{\text{Par}}_i^k$ is its value estimate provided by the GA via an estimated loop i , and L is the number of loops in a given model.

To describe specifically the errors made by UNet++ during the segmentation phase, two-dimensional cross-correlation coefficients (CCC) were calculated between the network-generated and the expected output distribution maps for each parameter:

$$\text{CCC}(\widehat{\mathbf{P}}_k, \mathbf{P}_k) = \frac{\sum_{\phi_j} \sum_{\Lambda_j} (\widehat{P}_{k\phi_j\Lambda_j} - \widehat{\overline{P}}_k)(P_{k\phi_j\Lambda_j} - \overline{P}_k)}{\sqrt{(\sum_{\phi_j} \sum_{\Lambda_j} (\widehat{P}_{k\phi_j\Lambda_j} - \widehat{\overline{P}}_k)^2)(\sum_i \sum_j (P_{k\phi_j\Lambda_j} - \overline{P}_k)^2)}}. \quad (30)$$

In Eq. (30), ϕ_j and Λ_j correspond to the latitudinal and longitudinal coordinate points of input and output maps $P_k(\phi, \Lambda)$ and $\widehat{P}_k(\phi, \Lambda)$ of each current loop parameter k .

Testing the methodology

Results using noise-free synthetic data

A relative freedom in the spatial arrangement of the sources which is the result of the unconstrained nature of the problem presented a significant difficulty to the estimation. A growth in reconstruction error with the

increasing number of current loops in the models was unavoidable. However, the refined neural net architecture and training method presented in Sect. "Refining the ML-based inversion" proved to be much more efficient at handling this challenge when compared to classical methods (see Appendix G) and the method presented in the introductory study. This quantitatively means that reconstruction errors started to grow significantly above a much higher number of sources in the models than in case of the original algorithm or using a more traditional least-squares inversion.

This trend can be observed in the graphs of Fig. 10 and for the NRMS and RPE metrics defined in Sect. "Loss functions and quality metrics". It is to be noted that these values were only derived for individual samples as it required a full computation including the GA estimation to be performed, which took considerable computational time.

As the new algorithm produced maps of reconstructed distributions for all the source parameters, cross-correlations (CCC calculated using (30)) between the true maps and neural net outputs were computed for subsets of multiple test samples containing the same number of current loops in the test set (see Table 8). These resulting cross-correlations and their average values were also plotted against an increasing number of loops in Fig. 11.

Perhaps the most apparent feature of the inference seen

in Fig. 11 is still the decrease of reconstruction accuracy against an increasing number of current loops in the model. Another noticeable tendency when looking at this graph is that the recovery of the rates of change in the loop currents presents a significantly more difficult task for the network than that of other source parameter distributions. The reason for that probably lies in the linear approximation of this parameter and its relatively complex relation to the finally formed magnetic fields.

In all the above graphs (Figs. 10 and 11) one can see the general trend, that errors start to grow significantly when over ~80 sources are admitted to the models, provided

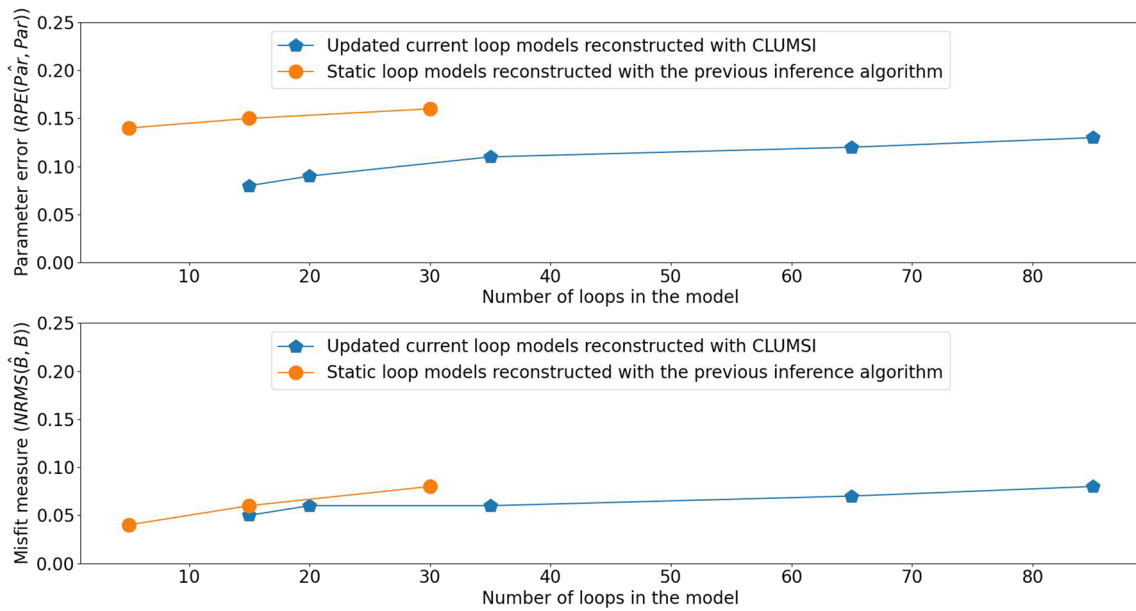


Fig. 10 RPE errors (a) and NRMS misfits (b) of individual final reconstructions carried out using the original and the refined (CLUMSI) algorithm against an increasing number of loops in the test data set

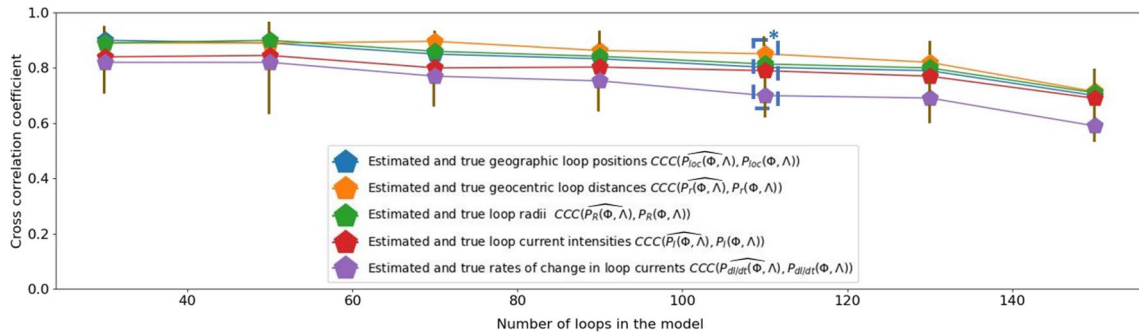


Fig. 11 CCC values between the reconstructed and true parameter maps plotted against an increasing number of sources in the corresponding models. These results were obtained using only the refined algorithm. Markers were assigned to the averages as shown on the label, the bars denote the min–max range of CCC values over a given subset. CCC values of the specific model featured also in Table 3 (being nearly the same as their corresponding subset averages) are referred to by the blue asterisk

that the refined machine learning framework is applied. This indicates a higher efficiency than that of the first framework, which already produced results with significant reconstruction error when applied for models containing only 15–20 sources.

The decrease in inference accuracy established itself as the more densely the sources were packed in the model domain, the harder it became even for the image-processing neural network to separate them based on their combined magnetic fields. This phenomenon is illustrated via

a comparison in Fig. 12, where the equivalence problem mentioned above is present in the outlined areas.

Broadly speaking, the test results suggest that reconstructions reproduce the general characteristics of regions with multiple sources in close proximity more reliably than the data coming from individual sources. Figure 12 suggests that overall characteristic features of the distribution are still recovered relatively well, even when the model contains a number of loops higher than 80. This is important to note because the model inferred

in case of real geomagnetic field data (Sect. “[Results on real geomagnetic model data](#)”) contained 116 current loops.

Graphs in Fig. 10a and b also show that while the new algorithm significantly outperforms the previous one in terms of inference error (RPE), it is still unable to do so when reconstructing the input magnetic field and SV maps (NRMS). To understand the reason behind this discrepancy, one needs to look at individual inversion results.

Figures 13 and 14 demonstrate an inference and field reconstruction of a synthetic model containing 30 current loops. It shows the distribution maps of the model parameters inferred with the help of the UNet++ algorithm along with maps of the radial field and SV values recovered via the final GA solution (see Sect. “[Refining the ML-based inversion](#)”). True maps are featured next to each corresponding reconstructed map for comparison.

When observing the reconstructed field maps in Fig. 14, one can notice that reconstruction quality is compromised by occasional outliers found mistakenly by the image-processing network. This kind of error is already somewhat observable on the ranges of CCC values in Fig. 11 as these mistakes can occasionally reduce the CCC of individual reconstructions thereby increasing the uncertainty of evaluations. Such erroneous estimations are highlighted in Fig. 13 showing the reconstructed loop parameter maps. It is a known general problem of applying deep learning algorithms for image segmentation, that though they can provide a fairly accurate overall inference, on some occasions they can produce very strong yet very wrong output signals (Popescu et al. 2021).

As the authors chose NRMS misfits of the SV values to account for only 10% in the total reconstruction loss (27) during the GA estimation process, unsurprisingly, the quality of the reconstruction of the radial magnetic induction is significantly better than that of the radial SV (see the maps and scales in Fig. 14c, d).

As for the stability of the inference with respect to potential noise in the input data, a first-order analysis is shown in Appendix F.

Results on real geomagnetic model data

In order to demonstrate the functionality of CLUMSI on actual geomagnetic field model data, in one reconstruction process, maps of the actual CMB radial magnetic induction field and SV at the 2019 epoch of the COVOBS geomagnetic field model (Gillet et al. 2019) were used as inputs. COVOBS data proved to be viable for a demonstration as they contain more relevant recent information

on the field, and they were relatively easy to use for constructing an input map (see Appendix D).

Figures 15 and 16 present the real and reconstructed geomagnetic field and SV maps. Based on the method described in Sect. “[Results using noise-free synthetic data](#)”, 116 current loops could be identified using the neural net output image. These are displayed as black dots on the map of the radial field. In addition, map c in Fig. 16 illustrates the relative importance of magnetic diffusion estimated by the reconstruction as the percentage ratio of the induced contribution and the total radial field in the GUFM-1 model (31):

$$W_{\text{diff}} = \frac{\widehat{B}_{r_{\text{tot}}} - \widehat{B}_{r_{\text{prim}}}}{\widehat{B}_{r_{\text{tot}}}}. \quad (31)$$

The overall quality of this reconstruction is similar to that of the synthetic results presented in Sect. “[Results using noise-free synthetic data](#)” in the sense that the radial induction vector field can be recovered with much higher accuracy than the radial SV and that occasional outliers in the segmented parameter maps significantly reduce the conformity between true and estimated fields. In general, however, it is ascertainable that the recovered current loop model does reproduce most of the main morphological features of the CMB radial field (and the SV as well for that matter, albeit much less accurately) provided by the geomagnetic model.

Figure 17 shows parameter maps of the model reconstructed using the input maps in Figs. 15 and 16. It is worthwhile to mention that interestingly, many of the reconstructed current loops seem to be distributed in chain-like arrangements, often aligned nearly meridionally (such occurrences are highlighted by dashed lines on the map of Fig. 17a)). It is however out of the scope of the present study to assess whether this can be linked to some property of the actual geodynamo or is an artifact of some sort. It is also noticeable that somewhat unexpectedly, the time variation of loop currents tends to show a more pronounced hemispherical dichotomy than the loop currents themselves (though we expect a much more uncertain recovery for these parameters, as it was established in Sect. “[Results using noise-free synthetic data](#)”).

Table 3 summarizes characteristic quality- and misfit parameters of the reconstructed field and the actual input fields. This too shows that the recovery of radial SV values and their distribution yielded a much poorer result than that of the actual radial field, manifesting most notably in a large difference between the corresponding CCC values.

Table 3 Misfit values of the reconstructed field compared to characteristic values of the corresponding geomagnetic field and SV magnitudes and the misfit values of a synthetic example for reference

	CMB radial field		CMB radial SV	
Quality parameters (misfit)	MAE	CCC	MAE	CCC
	0.17 [mT]	0.73	10.02 [mT/yr]	0.22
Characteristic values of the actual field	Absolute AVG	Absolute MAX	Absolute AVG	Absolute MAX
	0.264 [mT]	0.81 [mT]	10.19 [mT/yr]	70.5[mT/yr]
NRMS (misfit)	0.1		0.179	
NRMS (misfit) synthetic example with 110 loops (marked by * in Fig. 11)	0.075		0.17	

Discussion

To conclude this paper, it is important to underscore that the CLUMSI methodology presented in this study is still by no means perfect and comes with its own drawbacks.

On one hand, the most apparent limitation of the still highly idealized loop model aiming to represent the

current density distribution in the core is that it is unable to account for realistic SV magnitudes.

On the other hand, the most significant problem of the inversion framework comes from outliers in the neural net inference (Sect. “Results using noise-free synthetic data”) which in turn reduces the overall accuracy of reconstructions. Further complicating this issue is

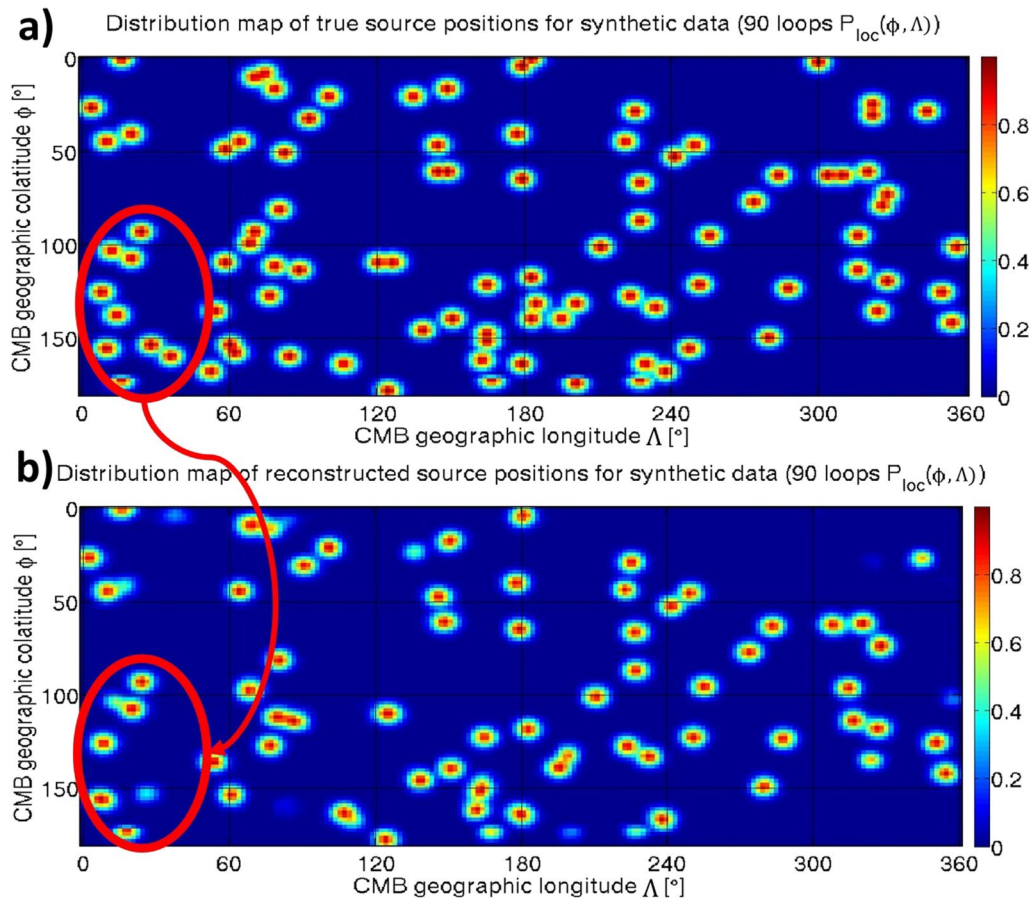


Fig. 12 Maps of true loop positions (a) and the positions estimated using the refined algorithm (b) for a model with 90 loops. One region is highlighted with a red circle where it is clearly visible that the neural network had difficulty in separating the magnetic signal of the loops

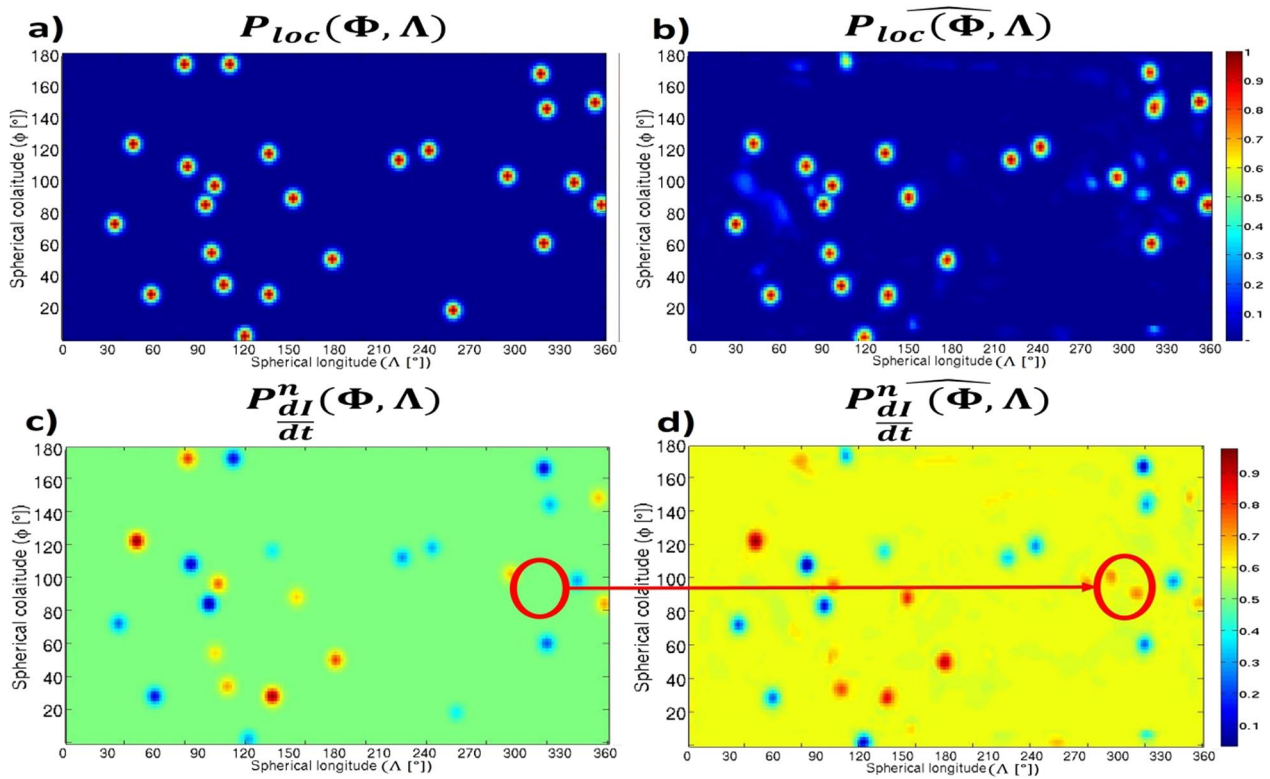


Fig. 13 True (left column) and reconstructed (right column) maps of the corresponding model parameters (60), (63) presented in Appendix C. Gross misidentifications are marked by the outlined areas

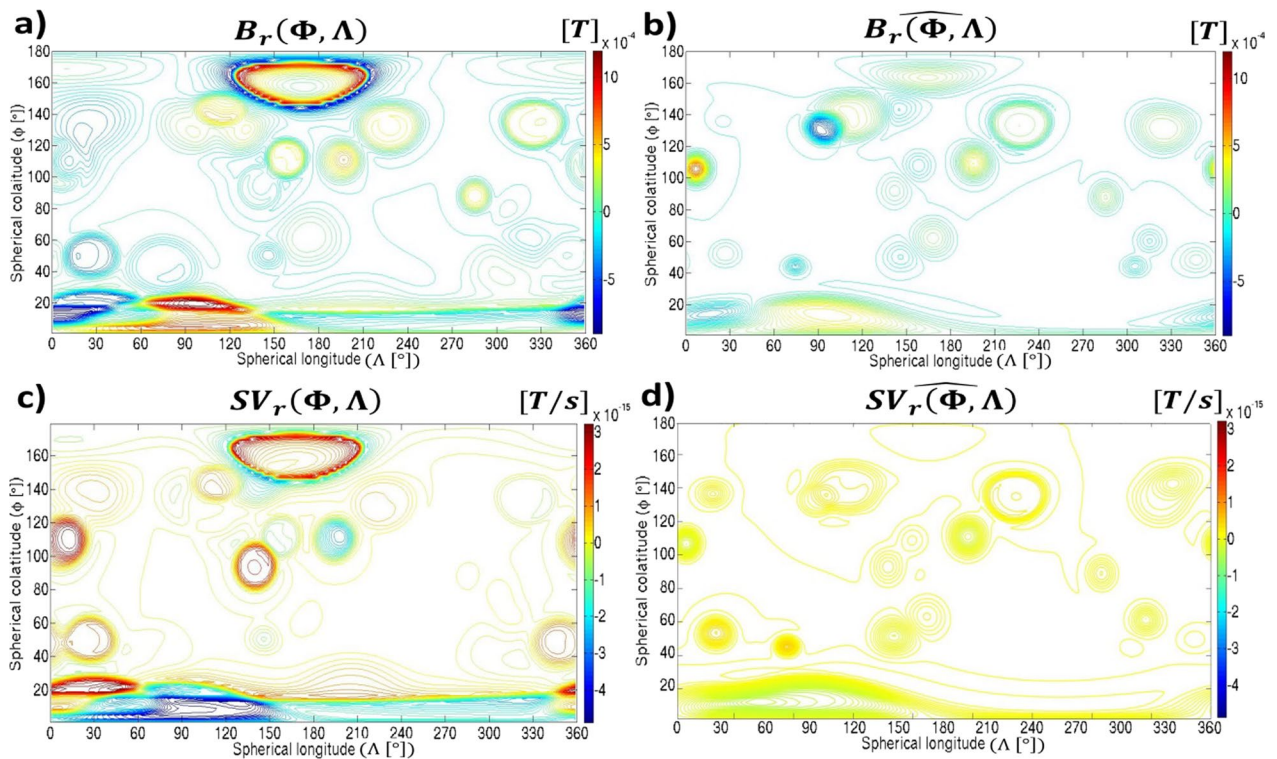


Fig. 14 Contour maps of the true and reconstructed magnetic fields and SV values in case of the same synthetic model as shown on Fig. 13

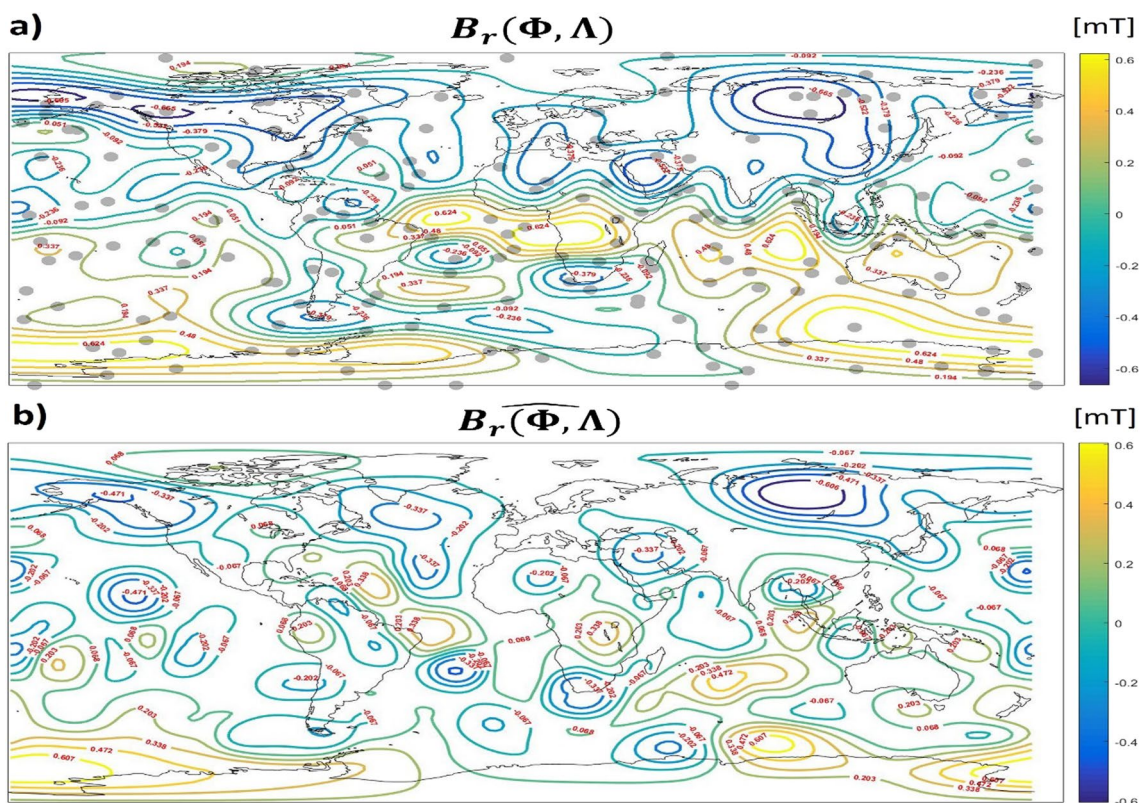


Fig. 15 True (a) and reconstructed (b) radial magnetic field values estimated using the 2019 epoch of the COVOBS geomagnetic field model at the CMB. The grey dots on map a indicate the source positions estimated using the network

that regularizations, such as physical constraints about inferred local field magnitudes or loop parameters, are not yet possible to directly integrate into the neural net training.

It is also important to note that the rectangular maps CLUMSI currently needs to rely on because of the type of neural network (UNet++) involved, cannot generate a truly uniform random distribution of loop currents on spherical surfaces. This can also introduce projection distortions when fitting a loop model to the observed field values.

These drawbacks can nevertheless be offset by the fact that CLUMSI can recover complex loop models more effectively (with a single evaluation and with higher accuracy) than previous techniques. Also, by the prospect that the approach has a very wide scope for further modifications and improvements with plenty of opportunities to enhance its physical authenticity and accuracy. Such ideas for further development are described below.

The reason for the loop models' inability to produce realistic SV is the lack of spatial movement of the features in the modeled fields. This issue can be resolved in the future by getting rid of the part of condition (1) (Sect. "Approximating the total field assuming linear time

variation in the source currents"), which imposes spatially fixed loops in the forward model. Possible SV values such forward models could produce would then be orders of magnitude higher at each surface location, than those which are currently achievable. Implementing this would however necessitate considering the effect of electromagnetic induction on the field of a moving current system. One possible way to computationally reproduce the effect could be the use of simplified 3D forward model samples applying a moving mesh finite element framework configuration. This additional complexity could be a reasonable extension of our existing loop model because it is known that a significant part of the locally registered SV comes from westward drifting features in the geomagnetic non-dipole field, which can move as fast as 0.5° in geographic longitude per year.

One avenue for alleviating the problem of outliers in the neural net inference could be to indirectly constrain and regularize the training and evaluation process via a custom loss function (see e.g., (Basir and Senocak 2022)). For that, the relation of the CLUMSI methodology to physics informed neural networks (PINN) and to inverse PINNs especially (Raissi et al. 2019; Jarolim et al. 2023) need to be explored. A larger training data set including

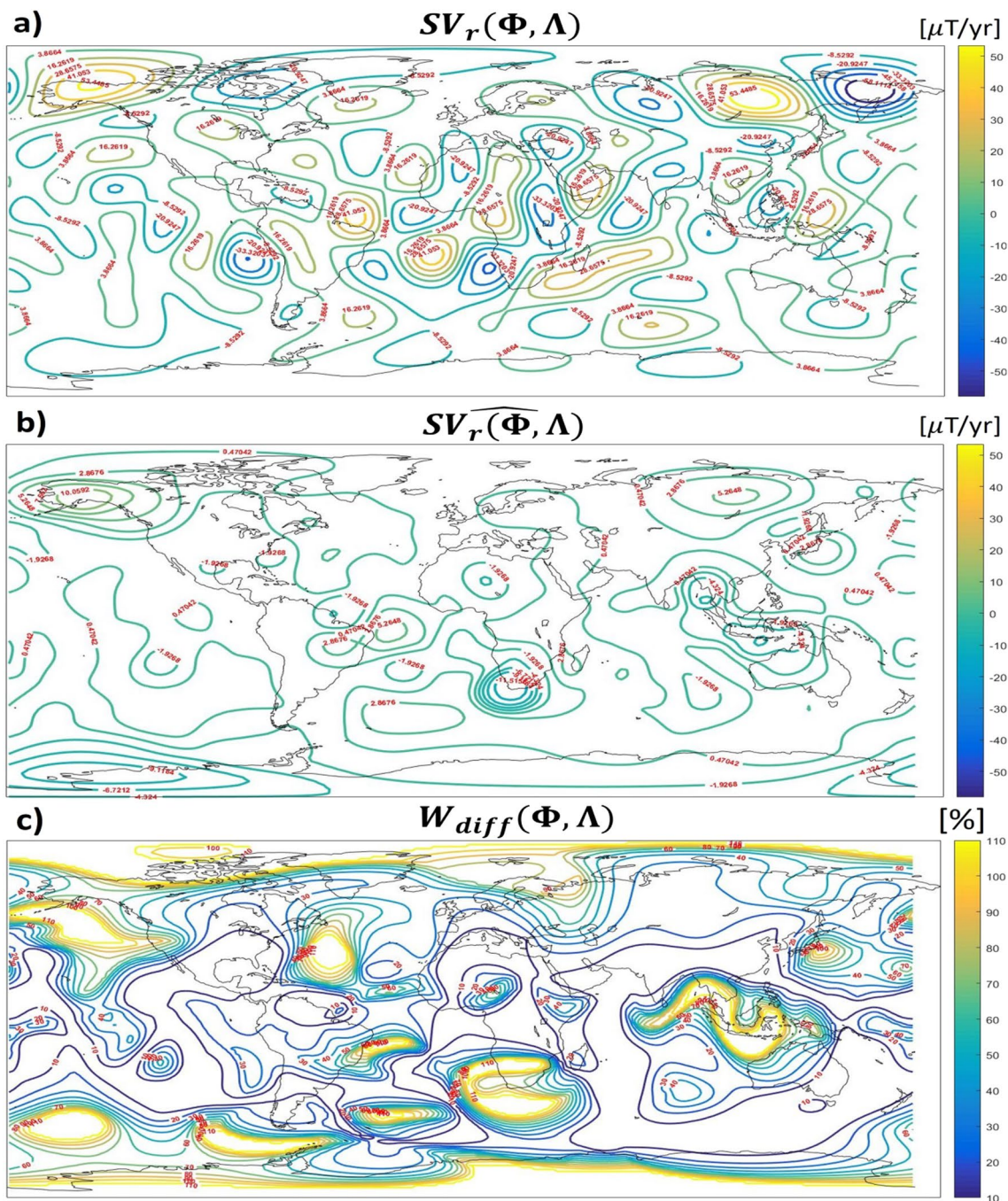


Fig. 16 True (a) and reconstructed (b) radial SV values estimated using the 2019 epoch of the COVOPS geomagnetic field model at the CMB. The relative importance of magnetic diffusion (c) in the reconstruction was illustrated using (31)

more synthetic and measured samples as well could also help in obtaining more accurate inferences. Concerning measured geomagnetic data, it is particularly important to note that the GUFM-1 model utilized in the training data set of this study makes use of historical observatory and maritime measurement data taken only at Earth's

surface (Jackson et al. 2000). These could be complemented by more recent geomagnetic model data, such as the COVOPS (Gillet et al. 2019) data, which are based on satellite measurements as well as surface observations, and could provide independent training data at multiple reference surfaces. This would potentially further

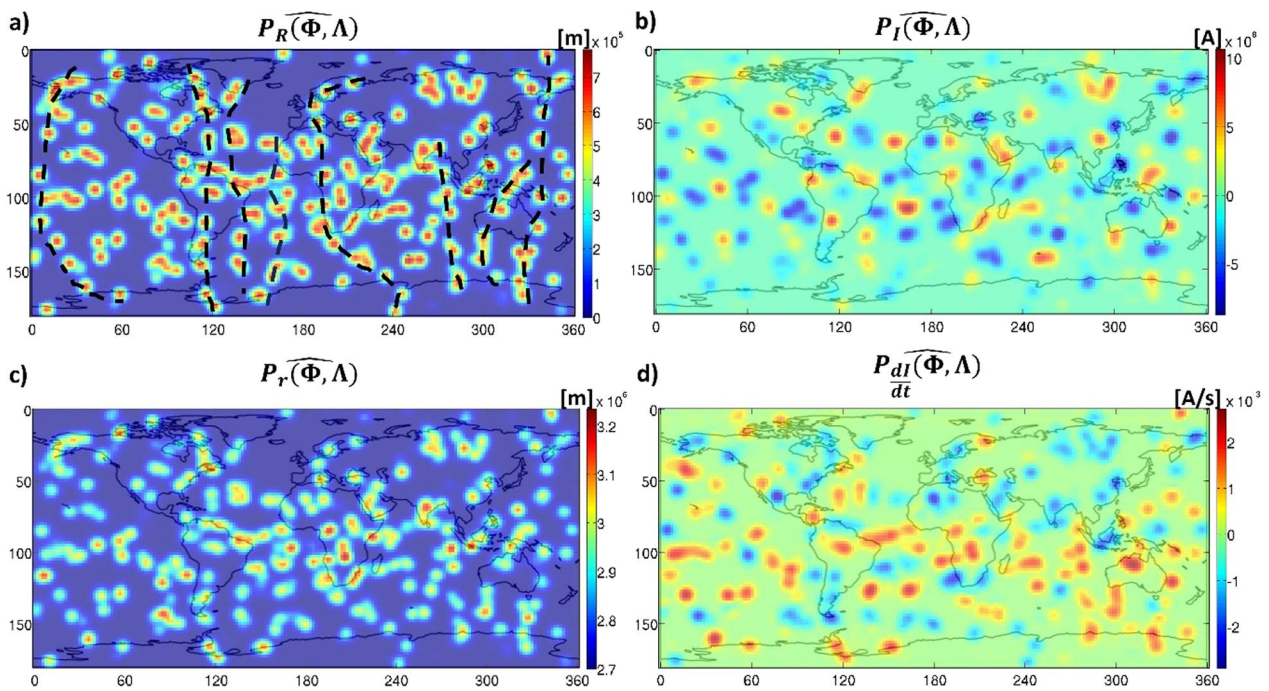


Fig. 17 Reconstructed maps of current loop parameters (60–63) segmented using actual geomagnetic field model data (COVOPS, epoch 2019)

improve the representativeness of the reconstructed source model with respect to the actual physical state of the geodynamo (Alken et al. 2021).

Current loop models inferred in this study tend to be local compared to the scale of the core. It remains an open issue if we can relate them to larger-scale current systems in the geodynamo and how this could be achieved. Further comparisons to MHD-based methodologies need to be conducted. For example, it could prove to be useful to compare snapshots of inferred loop parameter distributions, such as the ones presented in Figs. 15a and 17 taken using different epochs of an input geomagnetic model and compare them with maps of reconstructed core flows considered to dominate the intermittent time period.

It is also noteworthy that some derived parameter distributions in direct numerical simulations, like the “dynamo generation term” introduced by Miyagoshi et al. (2011) may display sufficient spatial stability and concentration while showing enough variability as well to be suitable for a similar kind of inference on real geomagnetic data.

Perhaps the future viability of CLUMSI (and similar attempts) as a supplementary tool for validating efforts

trying to picture the internal dynamics of the geodynamo depends on whether utilizing a composite of a high number of relatively simplistic sources can correctly represent such complex physical processes in general.

Conclusions

This study demonstrates the potential effectiveness of using deep learning for recovering highly complex current density distributions such as the one expected to be responsible for generating the magnetic field in the geodynamo. Synthetic tests confirmed that an image-processing neural network can recover complex distributions of source currents from input magnetic field data with reasonable accuracy. Main morphological features of the actual geomagnetic radial field could also be reproduced, albeit this was still based on a highly idealized equivalent current loop model and suffers from some significant drawbacks. Most notably, the inability to produce reconstructions that are representative of both the geomagnetic field and its SV, and occasional outliers in the inference which significantly affect the reconstruction quality.

Appendix A: Validation of the numerical model and detailed description of the computational framework

As a simple guiding value for setting the computational mesh size, the Courant criterion used in finite difference simulations (Dutykh 2016) was applied. This gives the following relation between time steps Δt and spatial discretization Δx for the classical magnetic diffusion problem:

$$\Delta t \leq \sigma \mu \Delta x^2 / 2. \quad (32)$$

Equation (32) for a minimum time step $\Delta t = 3.15 * 10^7$ [s] (1 year) yields a minimum finite element size of $\Delta x = 3200$ [m].

Equations (12–15) were solved for the vector potential A with an initial value:

$$A = 0. \quad (33)$$

Equation (33) applied for the entire domain as for estimating the induction screening discussed in Sect. “Approximating the total field assuming linear time variation in the source currents”, the current intensity could be linearly increased in time from $I_{t_0} = 0$ [A] at a rate given in Table 2.

The full solution used the MUMPS solver (Multifrontal Massively Parallel Solver 2022) on the linearized finite element discretization of Eqs. (12–15) using initial condition (33).

To validate whether the numerical models give a correct approximation of the induced magnetic fields, qualitative comparisons can be made for models which assume a conductive core. A process similar in nature to that described in Sect. “Approximating the total field assuming linear time variation in the source currents” results from the theoretical problem discussed in Weaver (1994), in which the cessation of the source current variation results in the restitution of the static primary field after the transients have ceased.

When assuming an insulating core, the fields obtained in the simulations can be checked by comparing them against the analytical solution presented in Sect. “Primary field of stationary current loops” on reference surfaces picked at different radial distances from the CMB. Comparisons of the solutions can be made for the magnetic field of a single source. In Fig. 18 resulting radial fields are plotted in the reference frame of one such current loop.

Figure 18 confirms that the two solutions, calculated on meridional circles with an increasing radius, agree to a high degree of accuracy in the order of magnitude range of the source parameters used for the training set.

Appendix B: Defining the range of model parameter values in the training set

The computation of synthetic training samples was performed using randomly generated values of the source parameters (see Table 1). To determine the range of possible magnitudes for these values, simple considerations concerning the physical dimensions of the core and the geomagnetic field were taken into account.

The CMB and the ICB arise as natural constraints on the spatial location of the reconstructed current loops. However, due to the high electrical conductivity of the core, the lower depth limit can be placed much higher. Let us imagine placing time-varying but spatially confined magnetic fields in a conductive core. In this case, a δ_z skin (or attenuation) depth can be given because of the electromagnetic screening of the conducting material above them (Gubbins 1996):

$$\delta_z = \left(\frac{\tau_p^l}{\mu_0 \sigma} \right)^{1/2}. \quad (34)$$

Equation (34) expresses this screening in spherical harmonic functions, where τ_p^l is the diffusion time associated with poloidal magnetic fields of degree l and order j . It is reasonable to assume that from time-varying local sources, most of the screened part of their magnetic fields (required for the separation of their signal) is essentially in the non-dipole SH spectrum. Approximating the core conductivity by $\sigma = 5 * 10^5$ [S/m] (see Sect. “Approximating the total field assuming linear time variation in the source currents”) and $\tau_p^{21} \approx 8000$ [yr] these fields can only emerge from within an upper layer of $\delta_{z_{\max}} \approx 800$ [km]. The upper limit for possible source depth in the models was therefore taken accordingly.

To give an initial estimate of the bounds on the possible values of the other source parameters, further considerations were made based on previous simulation results, statistical constraints on the geomagnetic field and the evolution of the geomagnetic dipole moment (GMD).

Gillet et al. (2010) provided a benchmark for the range of magnetic induction intensities within the Earth’s core. Their results show that, at least the radial component of the induction vector inside the Earth’s core can vary roughly between

$$B_{\min} \approx 1 \text{ [mT]} \quad (35)$$

and

$$B_{\max} \approx 6 \text{ [mT]}. \quad (36)$$

These values were used as limitations on the possible range of values for the interior field.

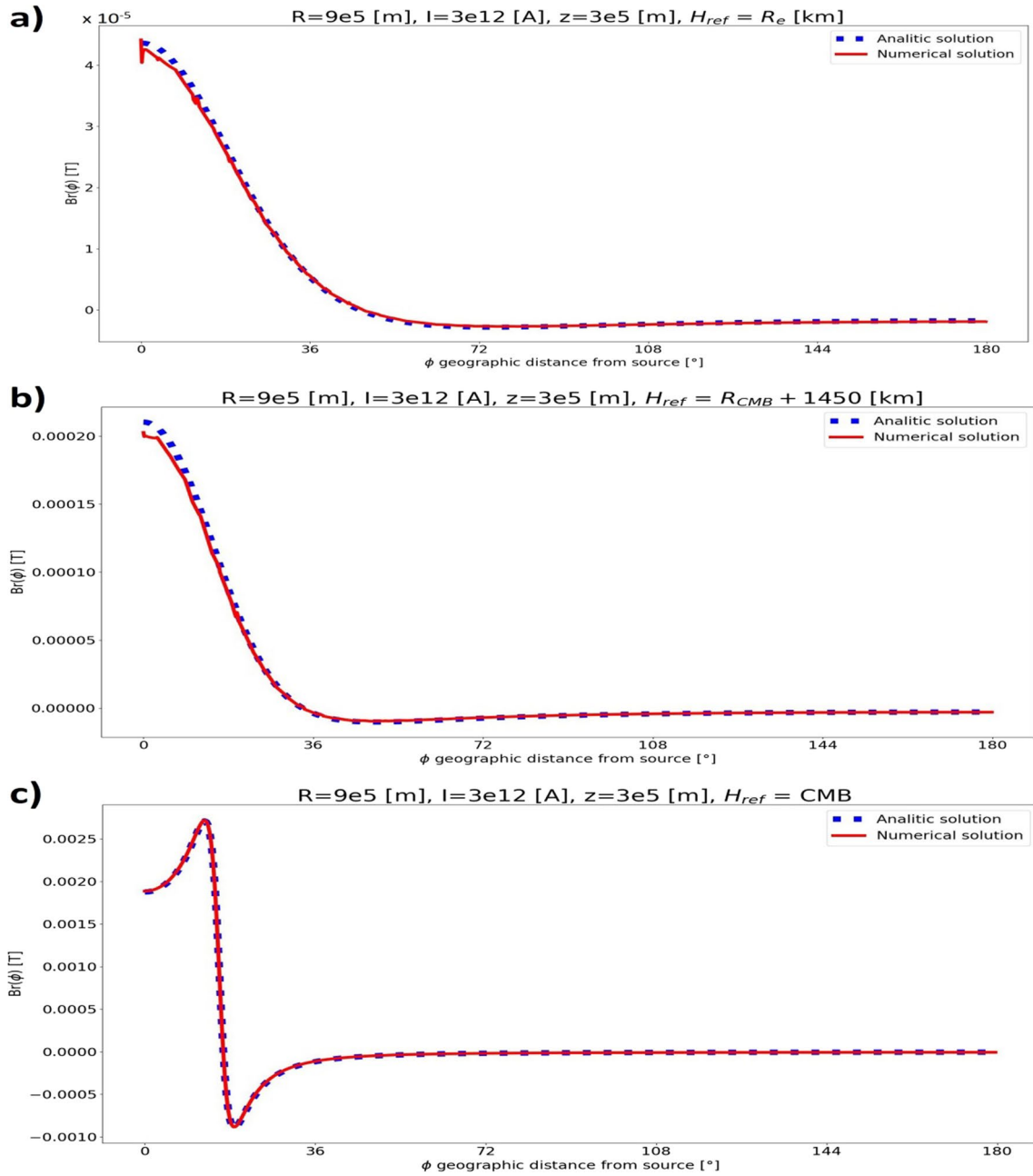


Fig. 18 Validating the numerical simulation against the analytical solution assuming an insulating core. The numbers above subplots **a**, **b** and **c** describe the source parameters and the radial position of the reference surfaces H_{ref} with respect to the CMB (R_e in graph **a**) is Earth’s radius)

Using this, six simplified independent relations (rules) between the extreme values of current loop parameters and the quantities GMD, B_{min} and B_{max} were constructed, represented by Eqs. (37–38), (39–40) and (41–42), respectively. The latter three quantities were considered to be known. In defining the relationships, sources were treated as current loops carrying constant currents and having fixed positions:

$$B_{min} = \mu I_{min} (2R_{max})^{-1}, \tag{37}$$

$$B_{max} = \mu I_{max} (2R_{min})^{-1}. \tag{38}$$

Equations (37) and (38) are related to combinations of source parameters that result in the largest and smallest absolute values of magnetic induction at the center of the loops (Fig. 19). The known minimum and maximum

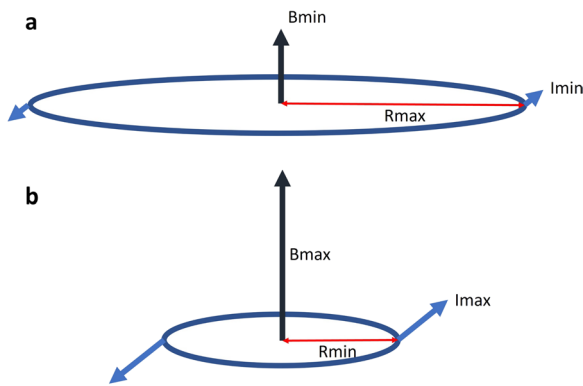


Fig. 19 Determination of the rule for B_{\min} (a) and B_{\max} (b) using circular currents taking extreme values of source parameters. R_{\min} and R_{\max} denote the maximum and minimum circular current radii sought, and I_{\min} and I_{\max} denote the maximum and minimum currents carried by the loops

magnitudes of the internal field cannot be exceeded in these cases, thus, values (35) and (36) can be assigned to their left-hand sides.

Two extreme configurations can also be considered. One in which the loops with the strongest current and the largest spatial extent give rise to the GMD together, and another in which the smallest loops carrying the weakest current do the same (Fig. 20). Then, Eqs. (39), (40) are simply given by aligning the axes of these current systems in these hypothetical models with the direction of the GMD.

Here, as the number of sources, N_{\max} is assigned to the model containing the smallest loops and N_{\min} to the model containing the largest ones:

$$GMD = I_{\max} R_{\max}^2 \pi N_{\min}, \tag{39}$$

$$GMD = I_{\min} R_{\min}^2 \pi N_{\max}. \tag{40}$$

Using two adjacent current systems in the two extreme models shown in Fig. 20, sharing a common axis, one can estimate how close these pairs can be to each other (see Fig. 21).

For this, field constraints (35–36) and the relation defining the axial magnetic field component along the axis of a current loop (Jackson 1998) can be exploited. On one hand, the distance of the two largest current loops furthest apart from each other (Fig. 21a) must be at most such that the minimum of the magnetic induction (halfway along the axis) cannot fall below constraint (35):

$$B_{\min} = 2 \frac{\mu_0 R_{\min}^2 I_{\max}}{2 \left((d_{\max}/2)^2 + R_{\min}^2 \right)^{3/2}}. \tag{41}$$

Similarly, the axial component of the magnetic induction along the common axis of a pair of the smallest current loops is at its maximum at the center of the loops. It can reach at most the value (36) in the model where the smallest current loops are closest to each other (Fig. 21b):

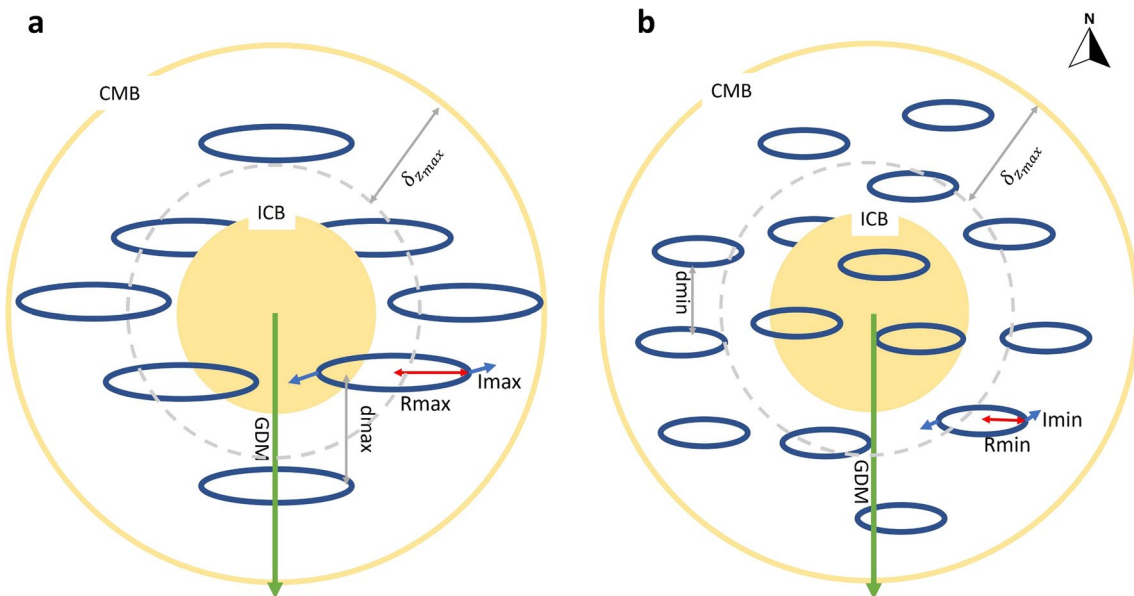


Fig. 20 Two extreme model configurations, in which the GMD is hypothetically produced by individual sources of the same size and elementary momentum which is in alignment with the GMD. $\delta_{z_{\max}}$ is the penetration depth corresponding to Eq. (34) (the figure is not to scale)

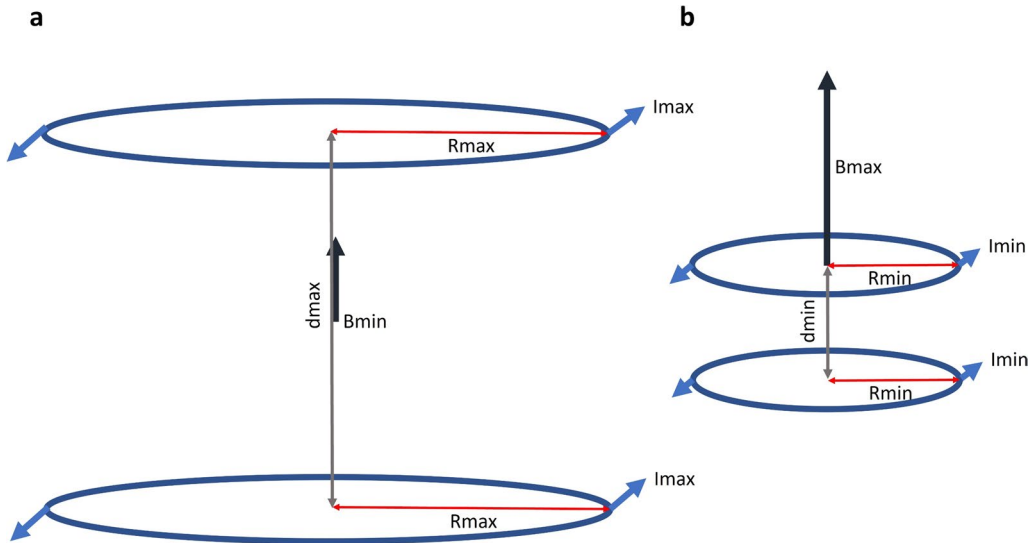


Fig. 21 The maximum d_{\max} (a) and minimum d_{\min} (b) distances allowed between the circular currents, belonging to the conceptual models shown in Fig. 20

$$B_{\max} = \frac{\mu_0 R_{\min}^2 I_{\min}}{2(d_{\min}^2 + R_{\min}^2)^{3/2}} + \frac{\mu_0 I_{\min}}{2R_{\min}}. \quad (42)$$

A further simplification is that the current loops in these models are uniformly distributed in the outer core. Their relative (equidistant) spacing is approximated by dividing the $\delta_{z_{\max}}$ thickness of the outer core volume into the same number of equal-sized cubes as the current loops in the models.

$$d_{\min} = \sqrt[3]{(4/3)\pi (r_{\text{CMB}}^3 - (r_{\text{CMB}} - \delta_{z_{\max}})^3) / N_{\max}} \quad (\approx 4.56 * 10^5 [\text{m}]), \quad (43)$$

$$d_{\max} = \sqrt[3]{(4/3)\pi (r_{\text{CMB}}^3 - (r_{\text{CMB}} - \delta_{z_{\max}})^3) / N_{\min}} \quad (\approx 1.57 * 10^6 [\text{m}]). \quad (44)$$

Equations (45) and (46) are then obtained by substituting Eqs. (43–44) into Eqs. (41–42).

In contrast to these hypothetical models, in forward solutions used for the reconstruction of the actual geomagnetic field, the direction of the axes of the current systems is radial based on other considerations (see Sect. “Approximating the total field assuming linear time variation in the source currents”). However, as it turns out (see Appendix C) these relations were suitable for the

Table 4 The minimum and maximum values of the model parameters, estimated using formulas (37–40) and (45–46)

	I [A]	R [km]	N
Minimum	$2.3 * 10^8$	325	25
Maximum	10^9	1000	1026

These provided the intervals between which the source parameters in the training samples were generated, as described in Appendix C

inclusion of parameter boundaries within which representative models of the CMB radial field could be generated, both qualitatively and quantitatively.

Since the resulting system of Eqs. (41–46) is non-linear, an iterative solution was chosen to solve them. This used an implementation of the Newton–Raphson method in MATLAB (Yang et al. 2005). The bounds defining Eqs. (41–46) were directly introduced by assigning the target values everywhere to the left and the variables to the right, which in case of (41–42):

$$B_{\min} = \mu R_{\max}^2 I_{\max} \left[R_{\max}^2 + N_{\min}^{-2/3} \left(2\pi \left(r_{\text{CMB}}^3 - (r_{\text{CMB}} - \delta_{z_{\max}})^3 \right) \right)^{1/3} \right]^{-3/2}, \quad (45)$$

$$B_{\max} = \mu I_{\min} (2R_{\min})^{-1} + \mu R_{\min}^2 I_{\min} \left[R_{\min}^2 + N_{\max}^{-2/3} \left(4\pi \left(r_{\text{CMB}}^3 - (r_{\text{CMB}} - \delta_{z_{\max}})^3 \right) \right)^{1/3} \right]^{-3/2}. \quad (46)$$

The final optimal solution (after ~1000 iteration steps) does not perfectly approximate any of the target values (see Table 5), however, even when applying the highest

Table 5 Validation against the characteristic target measures of the geomagnetic field when applying the derived limits of potential loop parameters in Eqs. (39)–(42) and Eqs. (45)–(46)

Target values	GMD = $8.1 * 10^{22} [\text{Am}^2]$	$B_{\min} = 10^{-3} [\text{T}]$	$B_{\max} = 6 * 10^{-3} [\text{T}]$
Values resulting from substituting the obtained maximum and minimum parameters into the referred conditions	(39) $\rightarrow \widehat{\text{GMD}} = 7.85 * 10^{22} [\text{Am}^2]$	(41) $\rightarrow \widehat{B}_{\min} = 1.5 * 10^{-4} [\text{T}]$	(42) $\rightarrow \widehat{B}_{\max} = 2 * 10^{-3} [\text{T}]$
	(40) $\rightarrow \widehat{\text{GMD}} = 7.84 * 10^{22} [\text{Am}^2]$	(45) $\rightarrow \widehat{B}_{\min} = 6.14 * 10^{-4} [\text{T}]$	(46) $\rightarrow \widehat{B}_{\max} = 6.2 * 10^{-3} [\text{T}]$

and smallest GMDs derived from GUFM-1 the training data (Jackson et al. 2000; Korte and Constable 2005), the solution does not change significantly (see Table 6).

Considering the rate of change in the source current over time, the following theoretical upper bound

$$(dI/dt)_{\max} = 1.13 [\text{A/s}]. \quad (48)$$

An important consequence of neglecting field diffusion here is described in Appendix C. Using the approximate solution introduced in Sect. “Approximating the total field assuming linear time variation in the source currents”, the total field of such a large change in the current above the largest sources could be as high as:

$$|B_{\text{tot}}^{\max}| = |B_{r_{\text{prim}}}^{\max} + B_{r_{\text{ind}}}^{\max}| = \left| \frac{\mu_0 R_{\max}^2 I_{\max}}{2(\delta_{z_{\max}} + R_{\max}^2)^{3/2}} + (dI/dt)_{\max} * C(R_{\max}, \delta_{z_{\max}}, 0) \right| \approx 0.074 [\text{T}] \gg F_u, \quad (49)$$

is obtained using the resulting R_{\min} from Table 4. If the change in the radial component of the induction vector directly above a source is investigated, and the effect of conductivity is neglected, a maximum (secular) change in the source current over time can be assigned to the maximum SV observed in the training data. If an axially aligned current loop in the model containing only the sources of the smallest spatial extent is considered (Fig. 20b), the local axial component of its associated magnetic induction vector coincides with the global radial component, from which it can be written (Jackson 1998):

$$\left(\frac{\partial I}{\partial t} \right)_{\max} = \frac{2(d_{\max}^2 + R_{\min}^2)^{3/2}}{\mu_0 R_{\min}^2} \left(\frac{\partial B_r}{\partial t} \right)_{\max}. \quad (47)$$

From the bounding Eq. (47) one can obtain that if $\left(\frac{\partial B_r}{\partial t} \right)_{\max}$ corresponds to the maximum radial SV component (measured on a 5-year basis) calculated from the GUFM-1 data set, then the current of any given current loop in the models can vary in time at most by

where $F_u \approx 0.0032 [\text{T}]$ is a value estimate based on the training data above which a field magnitude occurring in the generated samples can be considered an outlier (see Appendix C).

This means that source currents which have only a much smaller range in time variation can be applied when one aims at generating training samples representative in terms of actual geomagnetic field values. The simple procedure implemented to ensure this is described in Appendix C as well.

Table 5 contains the target values used for the computation of the range of possible current loop parameter values and their approximation using the derived extreme parameters. Table 6 represents the stability of the derived extreme parameters with respect to the changing GMD target values in the historical record.

Appendix C: Assembling and generating training and test data

In constructing the training dataset, synthetic magnetic field maps for the source models were generated within the bounds of source parameters discussed in Appendix

Table 6 Comparing the resulting lower and upper limits of potential loop parameters derived when using the weakest and strongest historical GMD magnitudes (Jackson et al. 2000) in the corresponding target values

Minimum and maximum historical magnitude of the GMD	$I_{\min} [\text{A}]$	$I_{\max} [\text{A}]$	$R_{\min} [\text{km}]$	$R_{\max} [\text{km}]$	N_{\min}	N_{\max}
$7.8 * 10^{22} [\text{Am}^2]$ (1975) ^a	$2.3 * 10^8$	10^9	324	1000	24	1025
$8.55 * 10^{22} [\text{Am}^2]$ (1830) ^a	$2.4 * 10^8$	10^9	335	1000	27	1027

^a denotes corresponding target values

B using the forward solution presented in Sect. “Approximating the total field assuming linear time variation in the source currents”.

Real geomagnetic data incorporated into the training samples were derived from the GUFM-1 historical geomagnetic field model (Jackson et al. 2000). The samples were Mercator maps of the CMB radial field with the same resolution as described in Sect. “Primary field of stationary current loops” taken from the model epochs using an increment value of 5 years.

A fundamental requirement for the synthetically generated part of any sample dataset used for ML is that it should contain values that are at least representative of the real data in terms of their order of magnitude (Huyen 2022). This was fulfilled by applying the sample selection criteria explained below.

A commonly used criterion for constructing a training data set is the selection of outliers, which can be defined using the quartiles of the distribution of the real data (in our case being a set A containing all radial magnetic field values coming from the GUFM-1 training data). This method is suitable also for data having non-normal distributions (Ilyas and Chu 2019; Czirok et al. 2022):

$$F_l = Q1(A) - 1.5(Q3(A) - Q1(A)) = -0.0031[T], \quad (50)$$

$$F_u = Q3(A) + 1.5(Q3(A) - Q1(A)) = 0.0032[T], \quad (51)$$

where $Q1$ and $Q3$ denote the first and third quartiles of the real data set (A).

Incorporation of a given sample into synthetic training data was rejected based on empirical conditions applied for field magnitudes which are detailed below.

The generation of an i individual current loop in the model was aborted if it did not satisfy the following three conditions within 100 repeated generation attempts.

1) The total field estimated directly above the source (estimated using Eq. (49)) fell within the range defined below:

$$\frac{F_l}{\log(N)} < B_{r_{tot}}^i < \frac{F_u}{\log(N)}. \quad (52)$$

2) The induced field of the source did not reduce the primary field to less than the primary field divided by e (i.e., the induction did not completely extinguish the primary field—a criterion related to Eq. (34) that gives the penetration depths for spherical diffusion times):

$$\frac{B_{r_{tot}}^i}{B_{r_{prim}}^i} > 1/e. \quad (53)$$

3) The generated source must fall further away from its nearest neighbor than an angular distance of 2° . This

ensured that, within the given resolution of the maps, the sources did not lie exactly underneath each other.

Conditions (1) to (3) can be checked already during the random generation of the individual source parameters, before the actual forward problem is solved, saving considerable computation time.

After solving the forward computation, the total magnetic field of each model sample was also checked using the condition below:

4) To ensure that the total field generated by the sample falls within the range defined by constraints F_l , F_u it was accepted if and only if,

$$F_l < \sum_i B_{r_{tot}}^i < F_u. \quad (54)$$

As it was postulated in Sects. “Approximating the total field assuming linear time variation in the source currents” and Appendix B, conditions (1)–(2) and (4) were chosen to be such that a training set representative in terms field magnitudes could be generated.

To alleviate the problem discussed in Appendix B concerning the rates of change in the current possible in the model samples, a systematic test was performed. It was conducted using a ‘damping factor’ γ applied on $\pm(dI/dt)_{\max}$, and summing up how many models can be accepted by varying source currents randomly within these values. Conditions (1) to (4) were applied as acceptance criteria. Figure 22 shows that for models with different numbers of sources below $\gamma = 1100$, the number of acceptable generated samples (models) starts to decrease in a similar fashion.

Finally, a training set containing 1030 Mercator maps generated using source models containing an incrementally increasing number of current loops (Table 7) and 75 maps coming from actual geomagnetic field values of the GUFM-1 model was assembled, and shuffled randomly when being loaded during the DANN training. Field maps in the GUFM-1 model were obtained directly at each geographical coordinate point given by the applied resolution (see Sect. “Primary field of stationary current loops”) and each epoch using the example program available in the `gufm1-webservice` (Rehfeld 2019). This was supplemented by maps of radial SV values readily available in case of GUFM-1 data on a 5-year basis and computed similarly using only the solutions for the primary fields in case of synthetic data (as a consequence of the approximation applied for the induced fields):

$$SV = (B_{r_{prim}}(I, t_0) - B_{r_{prim}}(I + dI/dt * \Delta t, t_0 + \Delta t)) / \Delta t, \quad (55)$$

where

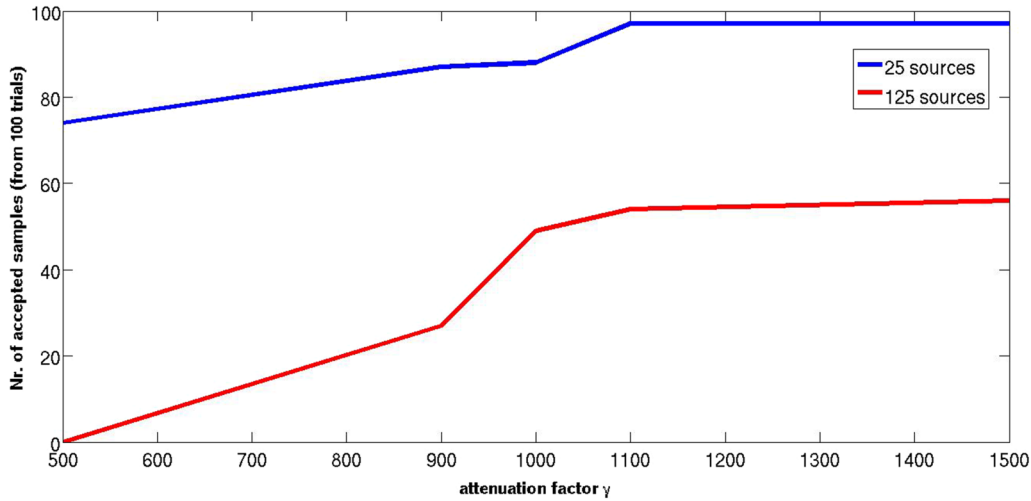


Fig. 22 Number of acceptable samples from 100 generated models, as a function of attenuation factor, for models with 25 and 125 sources

Table 7 Distribution of model samples in the synthetic training data set, containing 1030 Mercator (ϕ, Λ) maps of field and model parameter values

Number of loops in the model	25	35	45	55	65	75	85	95	105	115	125	135	145
Accepted model samples	100	100	100	100	100	100	100	99	94	79	46	11	1

$$\Delta t = 5 * 3.15 * 10^7 [s]. \quad (56)$$

Using this procedure synthetic magnetic fields could be generated which could produce radial magnetic components falling within the order-of-magnitude margins of the actual CMB radial field values derived from the GUFM-1 dataset. For comparison, histograms representing field distributions of the geomagnetic model and synthetic (loop) model values are shown on Figs. 23 and 24. In both cases, significant number of data points occur between $\pm 1[mT]$, however, the synthetic training data values have a near-normal distribution, whereas real training data have a depletion in a range of positive values due to the South Atlantic Anomaly (see e.g., (De Santis and Quamili 2010)).

Unsurprisingly, as the attenuation was applied for dI/dt in the current loop models, the bulk of synthetic data points are in an SV range 2 orders of magnitude

smaller than those coming from the GUFM-1 model (see Figs. 25 and 26).

To generate maps of target variables for the neural net training, normalized preprocessed maps of source parameter values were computed for each model as follows:

$$\mathbf{P}_k(\phi, \Lambda) = \mathbf{L}_k(\phi, \Lambda) p_k. \quad (57)$$

In Eq. (57), \mathbf{L}_k is a sparse matrix defined as

$$\mathbf{L}_k(\phi, \Lambda) = \begin{cases} 1 : (\phi, \Lambda) \equiv (\phi_{k_i}, \Lambda_{k_i}) \\ 0 : (\phi, \Lambda) \neq (\phi_{k_i}, \Lambda_{k_i}) \end{cases}, \quad (58)$$

$(\phi_{k_i}, \Lambda_{k_i})$ being the geographic coordinates of a given source i and p_k is a value of a given source parameter k for each corresponding current loop in the model. The preprocessing to optimize training performance is then carried out in the following fashion:

$$P_{kj}(\phi_{k_l}, \Lambda_{k_l}) = \max \left(P_{kj-1}(\phi_{k_l}, \Lambda_{k_l}), P_{kj-1}(\phi_{k_l}, \Lambda_{k_l}) e^{-\frac{(\phi_{k_l} - \phi_{k_i})^2 + (\Lambda_{k_l} - \Lambda_{k_i})^2}{2\sigma^2}} \right) \quad (59)$$

$$|k_l - k_i| \leq 5, k_l = 1, 2, \dots, 25.$$

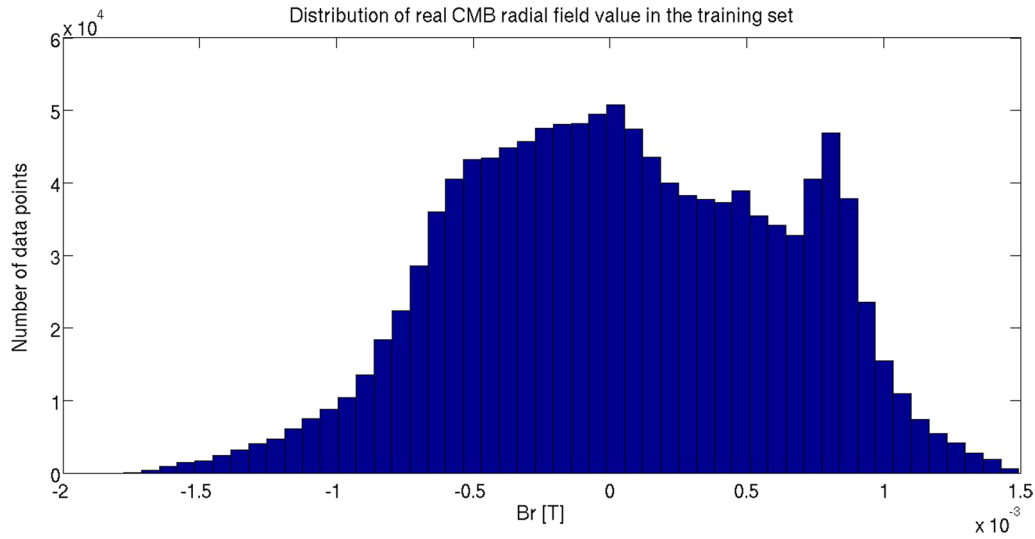


Fig. 23 Distribution of magnetic field values (radial component) on the CMB surface for the real geomagnetic models used for training (selected from the GUFM-1 model in 5 year epochs from 1600 to 1975)

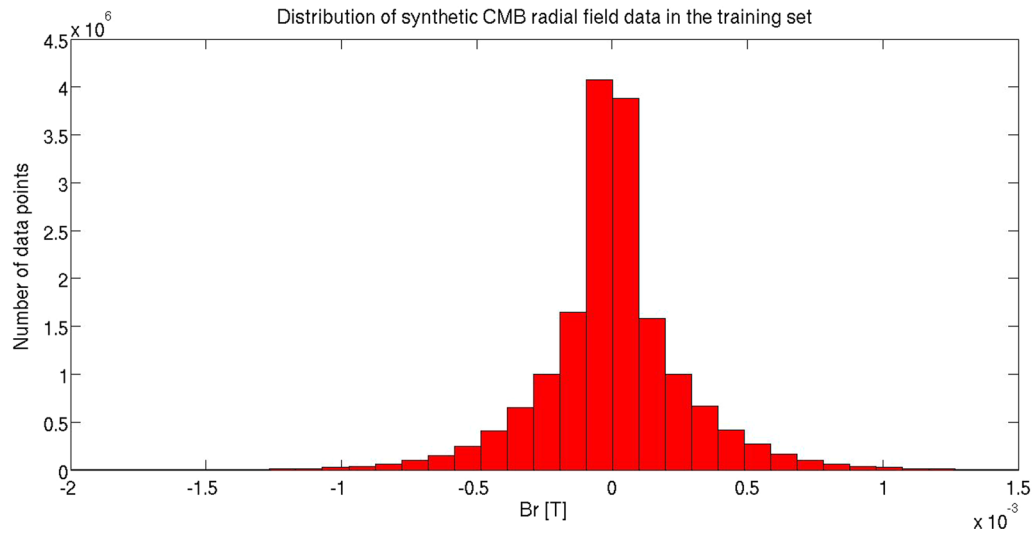


Fig. 24 Distribution of magnetic field (radial component) values on the CMB surface in all the models of the complete synthetic training set

In the above equation, $e^{-\frac{(\phi_{k_l} - \phi_{k_i})^2 + (\Lambda_{k_l} - \Lambda_{k_i})^2}{2\Sigma}}$ is a two-dimensional Gaussian kernel function, where Σ was set to 2 and j iterates through each grid point in a 5-by-5 wide window.

Parameter maps were then normalized using the maximum possible values of the source parameters defined in Appendix B:

$$P_r^n(\phi, \Lambda) = (P_r(\phi, \Lambda) - r_{\min})/\delta_{z_{\max}}, \tag{60}$$

$$P_{dl/dt}^n(\phi, \Lambda) = (P_{dl/dt}(\phi, \Lambda) + (dl/dt)_{\max})/(2 * (dl/dt)_{\max}), \tag{61}$$

$$P_R^n(\phi, \Lambda) = P_R(\phi, \Lambda)/R_{\max}, \tag{62}$$

$$P_I^n(\phi, \Lambda) = (P_I(\phi, \Lambda) + I_{\max})/(2 * I_{\max}). \tag{63}$$

Examples of the final maps containing real and synthetic data incorporated in the training set are shown in Figs. 27 and 28.

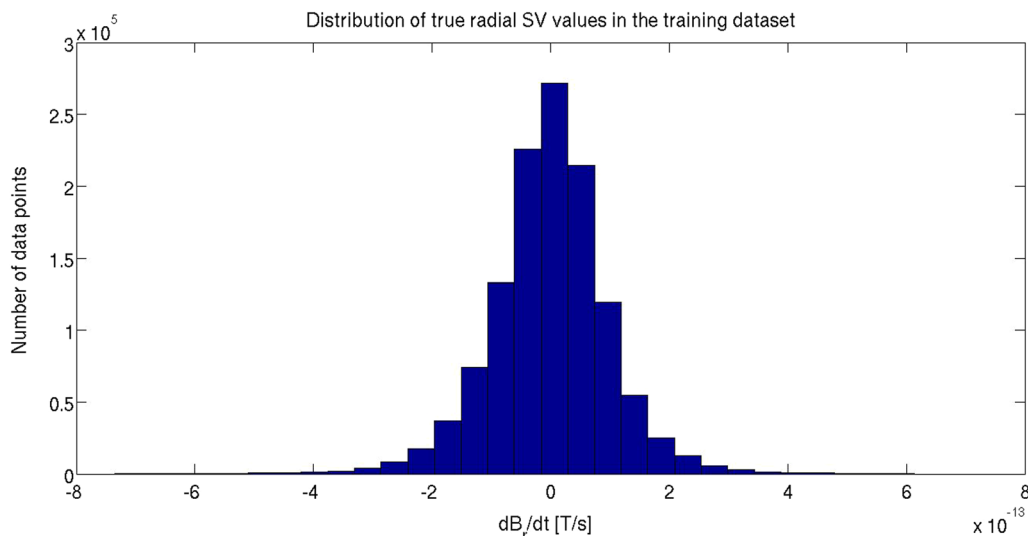


Fig. 25 Estimated distribution of SV values in the GUFM-1 data set used for the training (calculated on a 5-year basis)

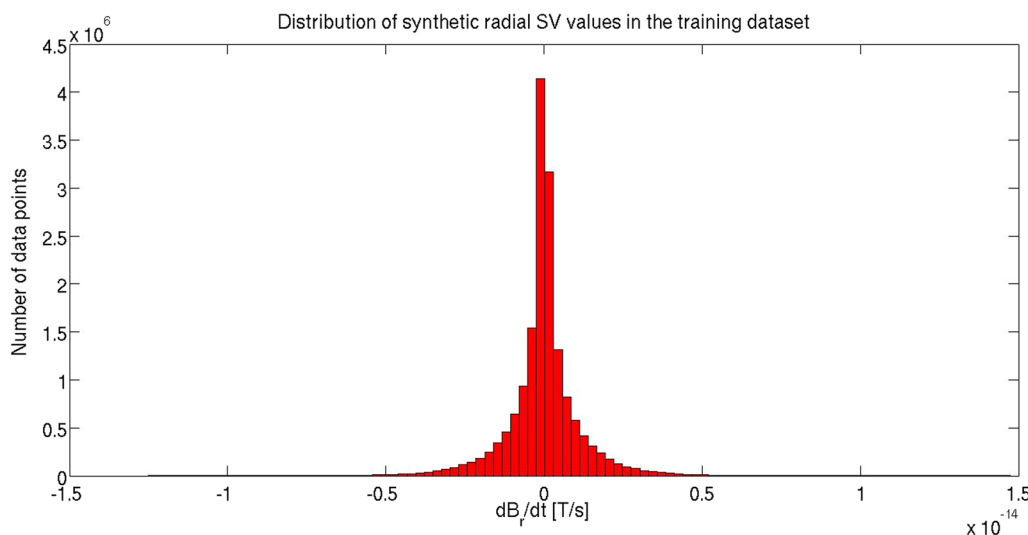


Fig. 26 Estimated distribution of SV values in the full synthetic data set (5-year basis). Note that in this figure, the x axis represents a range almost two orders-of-magnitude smaller range than in Fig. 25

Test data were also generated in the same fashion as the training data, but a smaller number of models was admitted, and test samples contained a differing number of loops than those of the training data set (see Table 8)

Appendix D: Deriving magnetic field and SV maps from the COVOBS geomagnetic field model

CMB radial field and secular variation data used for the demonstration in Sect. “Results on real geomagnetic model data” were extracted from the COVOBS

geomagnetic field model with the help of the pygeodyn python package (Huder et al. 2019). Maps were constructed using an SH expansion of the model coefficients up to degree 14 in pyshtools (Wieczorek and Meschede 2018).

Appendix E: Description of the genetic algorithm used for final reconstruction

The overall workflow of the GA remained the same as described in (Kuslits et al. 2020). However, the parameter vector:

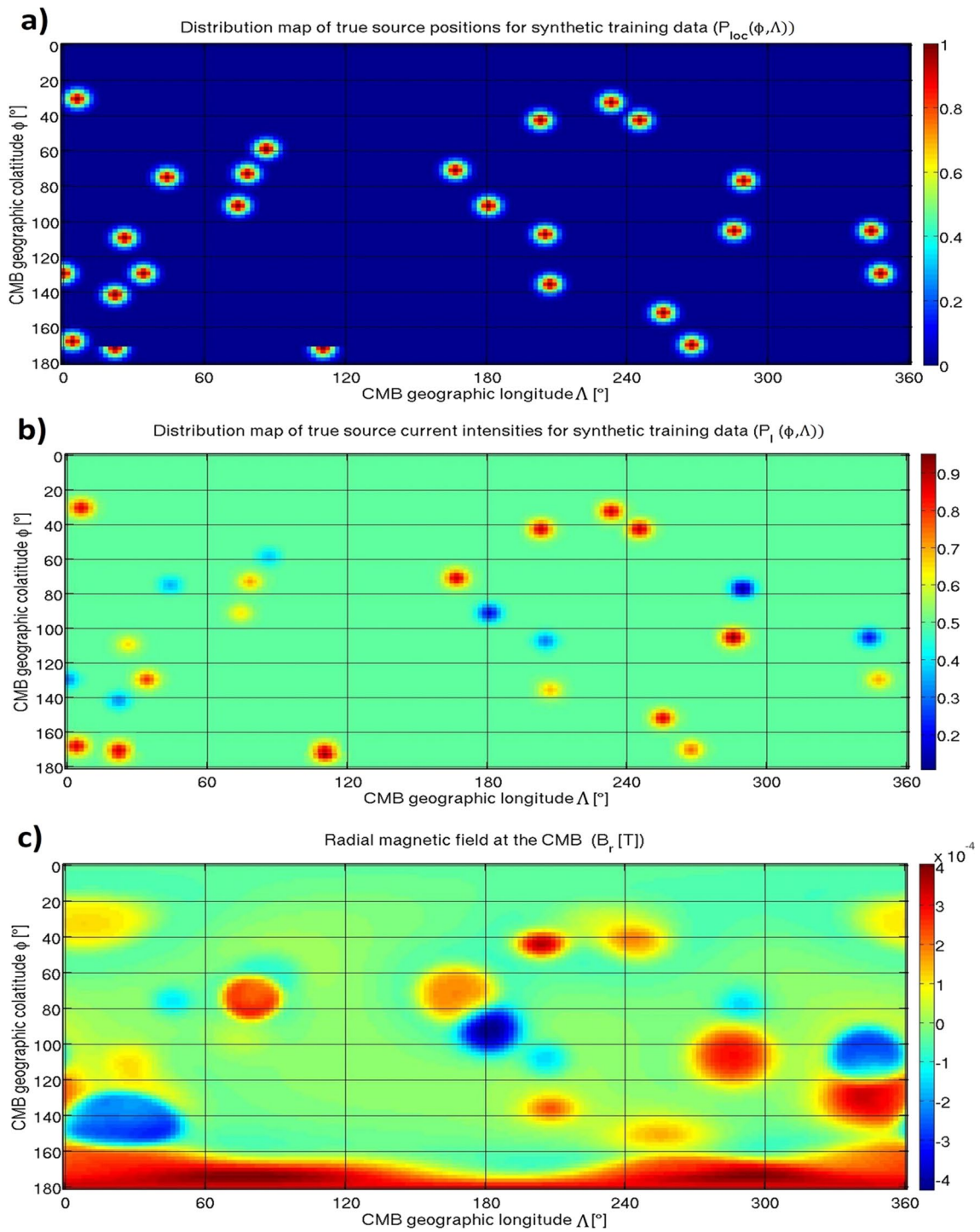


Fig. 27 Mercator maps of $P_r^l(\phi, \Lambda)$ (a), $P_I^l(\phi, \Lambda)$ (b) and radial magnetic field (c) in the synthetic component of the training set

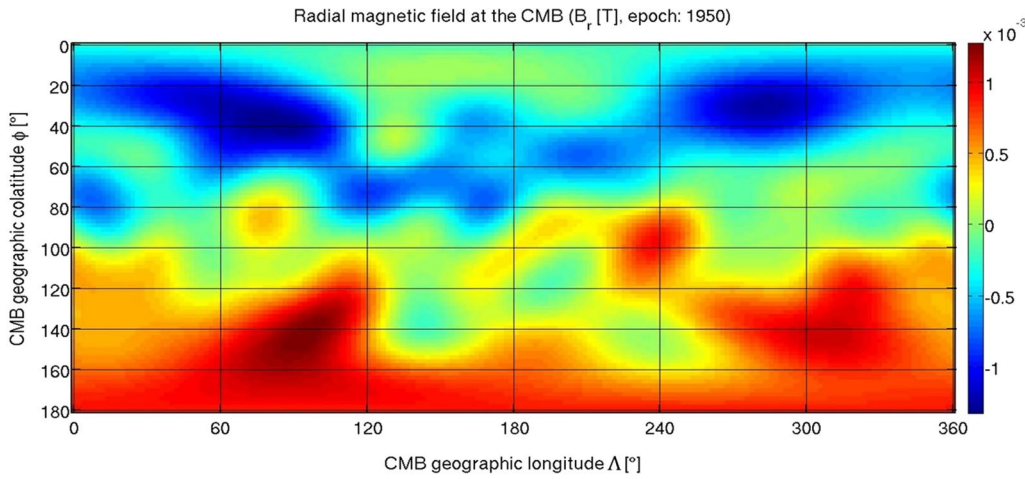


Fig. 28 Mercator map of the CMB radial magnetic field in the real component of the training set (GUFM-1, epoch 1950)

$$\mathbf{Par}_i = \begin{bmatrix} (\phi_i) \\ (\Lambda_i) \\ (r_i) \\ (R_i) \\ (I_i) \\ (\theta_i) \\ (\lambda_i) \end{bmatrix} \quad (64)$$

could be further constrained via sampling possible values for the elements of vector (64) from the UNet++ solution maps within 2-pixel intervals of (ϕ_i, Λ_i) (see Fig. 8), and applying a logarithmic transformation $\mathbf{Par}_i = [\log(\phi_i), \log(\Lambda_i)]$.

Initially, $N_{\text{ind}} * N_{\text{pop}}$ solutions are generated with randomly chosen parameter values, where N_{pop} is the number of populations and N_{ind} is the number of individuals in the population.

At the beginning of the selection phase, the program calculates for each specimen the deviation of the solution from the target value according to Eq. (27). Then, the target values are ranked and linearly transformed (fitness value) to select individuals, the better the fit, the higher the probability (Baker 1985):

$$\text{fitness}(\text{Pos}) = 2 - SP + 2(SP - 1)(\text{Pos} - 1)/(N_{\text{ind}} - 1). \quad (65)$$

In Eq. (65), Pos represents the position of a given specimen (physical model) when ranking of all the models according to (27) produced in a given step within the population. The parameter SP also called selection pressure was assigned a value of $SP = 2$. This being a real coded algorithm, parameters of the individual solutions

(\mathbf{Par}^{tr}) varied in a given population during crossover or recombination obeying Eq. (66) (Picek et al. 2013):

$$\begin{aligned} & (\mathbf{Par}(N_{\text{ind}_l}, \text{gen} + 1) | pp_{\text{rek}} = 1) \\ &= \chi \mathbf{Par}(N_{\text{ind}_{pt1}}, \text{gen}) \\ &+ (1 - \xi) \mathbf{Par}(N_{\text{ind}_{pt2}}, \text{gen}), \end{aligned} \quad (66)$$

where N_{ind_l} , $N_{\text{ind}_{pt1}}$, $N_{\text{ind}_{pt2}}$ are random numbers between 1 and N_{ind} within any given population, gen denotes the generation (number of iteration steps), and ξ is also generated as a (pseudo-) random number in the interval [0, 1]. Crossover probability between the selected individuals was set to $p(pp_{\text{rek}} = 1) = 0.6$. In the mutation step, the parameters of an individual could be randomly modified with a probability similar to the mutation rate for real coded algorithms in the work of Mühlenbein and Schlierkamp-Voosen (1993):

$$\begin{aligned} & (\mathbf{Par}(N_{\text{ind}_{\text{mut}}}, \text{gen} + 1) | pp_{\text{mut}} = 1) \\ &= \mathbf{Par}(N_{\text{ind}_{\text{mut}}}, \text{gen} + 1) \\ &+ \beta_1 (\mathbf{Par}_{\text{max}}(N_{\text{ind}})) \\ &- \mathbf{Par}_{\text{min}}(N_{\text{ind}}) \sum_{j=1}^a \beta_2 \beta_3^{-j}. \end{aligned} \quad (67)$$

In Eq. (67), the mutation occurred with probability $p(pp_{\text{mut}} = 1) = 0.02$, β_1 is a random number between -1 and 1 , β_2 is a sequence of random numbers with length a and elements $p(\beta_2 = 1) = 1/a$, $\beta_3 = 1.4$ (own setting) and $a = 20$ determines the smallest possible magnitude of the mutation effect. $\mathbf{Par}_{\text{max}}(N_{\text{ind}})$ and $\mathbf{Par}_{\text{min}}(N_{\text{ind}})$ denote the extreme values of each parameter, respectively, as defined in Appendix B. It is

Table 8 Distribution of model samples in the synthetic test data set

Number of loops in the model	30	50	70	90	110	130	150
Accepted model samples	50	50	50	50	50	15	7

important to note, however, that the effects of the steps listed above have always been limited to these extrema:

$$\text{IF : } \mathbf{Par}(N_{\text{ind}}, \text{gen}) > \mathbf{Par}_{\text{max}}(N_{\text{ind}}) \text{ THEN : } \mathbf{Par}(N_{\text{ind}}, \text{gen}) = \mathbf{Par}_{\text{max}}(N_{\text{ind}}) \quad (68)$$

$$\text{IF : } \mathbf{Par}(N_{\text{ind}}, \text{gen}) < \mathbf{Par}_{\text{min}}(N_{\text{ind}}) \text{ THEN : } \mathbf{Par}(N_{\text{ind}}, \text{gen}) = \mathbf{Par}_{\text{min}}(N_{\text{ind}}). \quad (69)$$

The set of individuals that underwent the mutation process formed the initial populations for the next iteration step. Individuals were also swapped between each other at a rate of 4% of the population size after a specified number of consecutive steps. This algorithmic logic, in which the evolution process of populations is run separately in parallel up to a specified iteration step interval is called a multi-population genetic algorithm (Grefenstette 1981). Exchanging some individuals among the populations at a given proportion is referred to as migration.

Achieving a globally optimal solution in case of synthetic data typically required several hundred to a thousand iteration steps. The reconstruction of real geomagnetic field and SV data presented in Sect. “Results on real geomagnetic model data” took 560 steps.

Appendix F: Results using noisy synthetic data

To perform a first-order stability analysis, noise was indirectly introduced to the synthetic magnetic fields via the perturbation of loop parameters normalized using the same corresponding quantities as in Eqs. (60)–(63). Perturbed normalized parameters \mathbf{Par}_N^n came from a normal distribution with an average \mathbf{Par}^n and standard deviation Σ :

$$\delta \sim |\mathbf{Par}^n - \mathbf{Par}_N^n|, \mathbf{Par}_N^n \sim N(\mathbf{Par}^n, \Sigma), \quad (70)$$

$$\widetilde{\mathbf{Par}}^n = \mathbf{Par}^n + \delta, \widetilde{\mathbf{Par}}^{n'} = \mathbf{Par}^n - \delta, \quad (71)$$

and after transforming back to dimensional values:

$$\widetilde{B}_{r_{\text{tot}}} = \frac{\widetilde{B}_{r_{\text{tot}}}(\widetilde{\mathbf{Par}}) + \widetilde{B}_{r_{\text{tot}}}(\widetilde{\mathbf{Par}}')}{2}, \quad (72)$$

$$\widetilde{SV}_r = \frac{(\widetilde{B}_{r_{\text{tot}}}(t_0 + \Delta t) - \widetilde{B}_{r_{\text{tot}}}(t_0))}{\Delta t}, \quad (73)$$

Δt being the 5-year basis mentioned in Appendix C.

This allowed for producing averaged noisy magnetic field and SV maps for a reference model with a given number of loops. The response of the full reconstruction of the physical model to the noise introduced to input data has so far only been investigated for a model containing 15 loops, as the necessary series of calculations could be performed in a relatively short time.

Figure 29 summarizes the results obtained in terms of both the relative error of the parameter estimation and the misfit (see Sect. “Loss functions and quality metrics”). It can be observed that the increase in noise level does not degrade the fit of the reconstructed magnetic fields, but significantly degrades the accuracy of the reconstruction of the loop parameters.

For models with a larger number of circular currents, a stability analysis of the image-processing phase and the reconstructed maps of source parameter distributions was performed.

For a model with 25 loops, Fig. 30 shows how much of the source distribution the image-processing network was able to reconstruct using noise-free and noisy synthetic inputs. Row c) of this tiled figure demonstrates how, with an increasing noise level, the image-processing

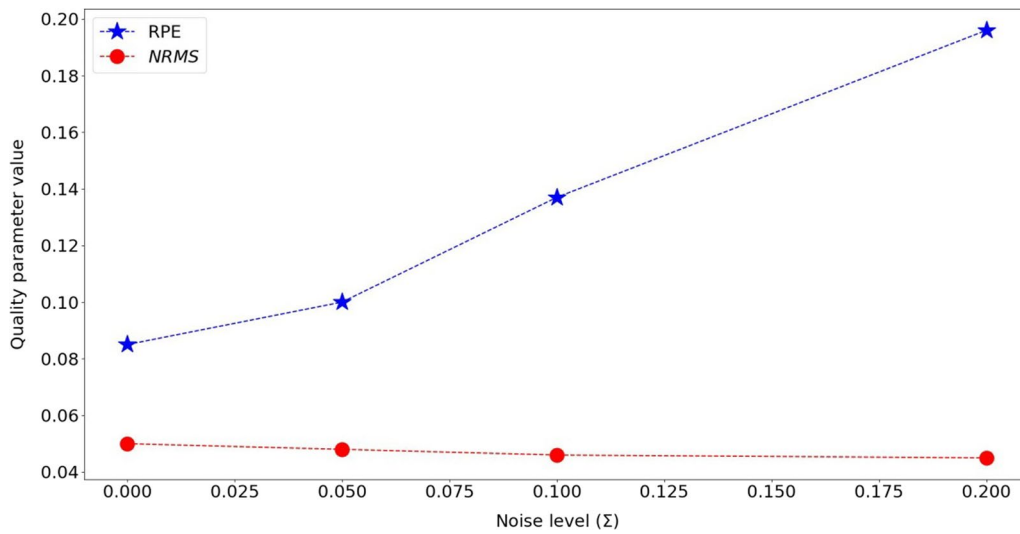


Fig. 29 Quality parameters $\overline{\text{NRMS}}$ (27) and RPE (29) plotted against an increasing noise level in the input data

network starts to infer more sources in the vicinity of loops that are actually there.

Figure 31 summarizes the cross-correlation values between the reconstructed and the true distribution maps. On one hand, the figure shows that even with the above-mentioned increase in error, the reconstruction of the geographical distribution remains in general the least sensitive to noise constructed using Eqs. (70–73).

On the other hand, with the inclusion of a larger number of sources to be reconstructed, the stability deteriorates significantly (although it should be noted that since the noise (70–73) is added separately to each loop parameter, it may be cumulative for a larger number of sources).

Appendix G: Difficulties with using a “classical” approach

One can legitimately think about the question of whether apart from the possibility of using the DANN methodology, the use of machine learning for the estimation task presented in this paper is justified at all. To demonstrate its advantages when compared to previous efforts (reviewed in Sect. “Introduction”) resorting to linear and non-linear inversion directly for recovering current loops, the authors reimplemented the rather ingenious attempt of Alldredge (1987).

His method was particularly interesting for the authors as though it did not take into account SV and magnetic diffusion, it too applied little a priori constraints besides assuming radially aligned current loops. It arguably uses more straightforward computations as it operates mostly in the spherical harmonic spectral domain and

offers good convergence as far as RMS misfit values are concerned.

The forward computation results in the spherical harmonic coefficients (SHC) representing a set of L loops using Eq. (74) described in detail in Alldredge (1987):

$$\widehat{g}_n^m = \sum_{l=1}^L K_l \frac{2n}{\sin(\alpha_l)} \left(\frac{r_l}{a}\right)^{n-1} \frac{P_n^1(\cos(\alpha_l))}{\sqrt{2n(n+1)}} P_n^m(\cos(\phi_l)) \cos(m\Lambda_l), \tag{74}$$

$$\widehat{h}_n^m = \sum_{l=1}^L K_l \frac{2n}{\sin(\alpha_l)} \left(\frac{r_l}{a}\right)^{n-1} \frac{P_n^1(\cos(\alpha_l))}{\sqrt{2n(n+1)}} P_n^m(\cos(\phi_l)) \sin(m\Lambda_l), \tag{75}$$

where r_l is the distance between Earth’s center and a given circular current loop l , α_l , is half of the viewing angle of a loop from Earth’s center, ϕ_l and Λ_l are the loop’s geographic longitude and co-latitude, and

$$K_l = \frac{10^9 \mu_0 r_l^2 \sin(\alpha_l) I_l}{a^3}. \tag{76}$$

The estimation of the loop parameters according to the definition used by Alldredge, $\mathbf{Par}_{Al} = [K_l, \alpha_l, r_l, \phi_l, \Lambda_l]$ was performed via a least-squares fitting of the SHC:

$$\widehat{\mathbf{Par}}_{Al} = \operatorname{argmin} \left(\left(\widehat{g}_n^m(\widehat{\mathbf{Par}}_{Al}) - g_n^m \right)^2 + \left(\widehat{h}_n^m(\widehat{\mathbf{Par}}_{Al}) - h_n^m \right)^2 \right). \tag{77}$$

The standard linear inversion was applied for K_l , and a variant of the Marquardt–Levenberg algorithm for the rest of the loop parameters which are in a non-linear relationship with coefficients g_n^m, h_n^m .

It turns out however, that when one tries to practically implement and test this technique, a handful of issues

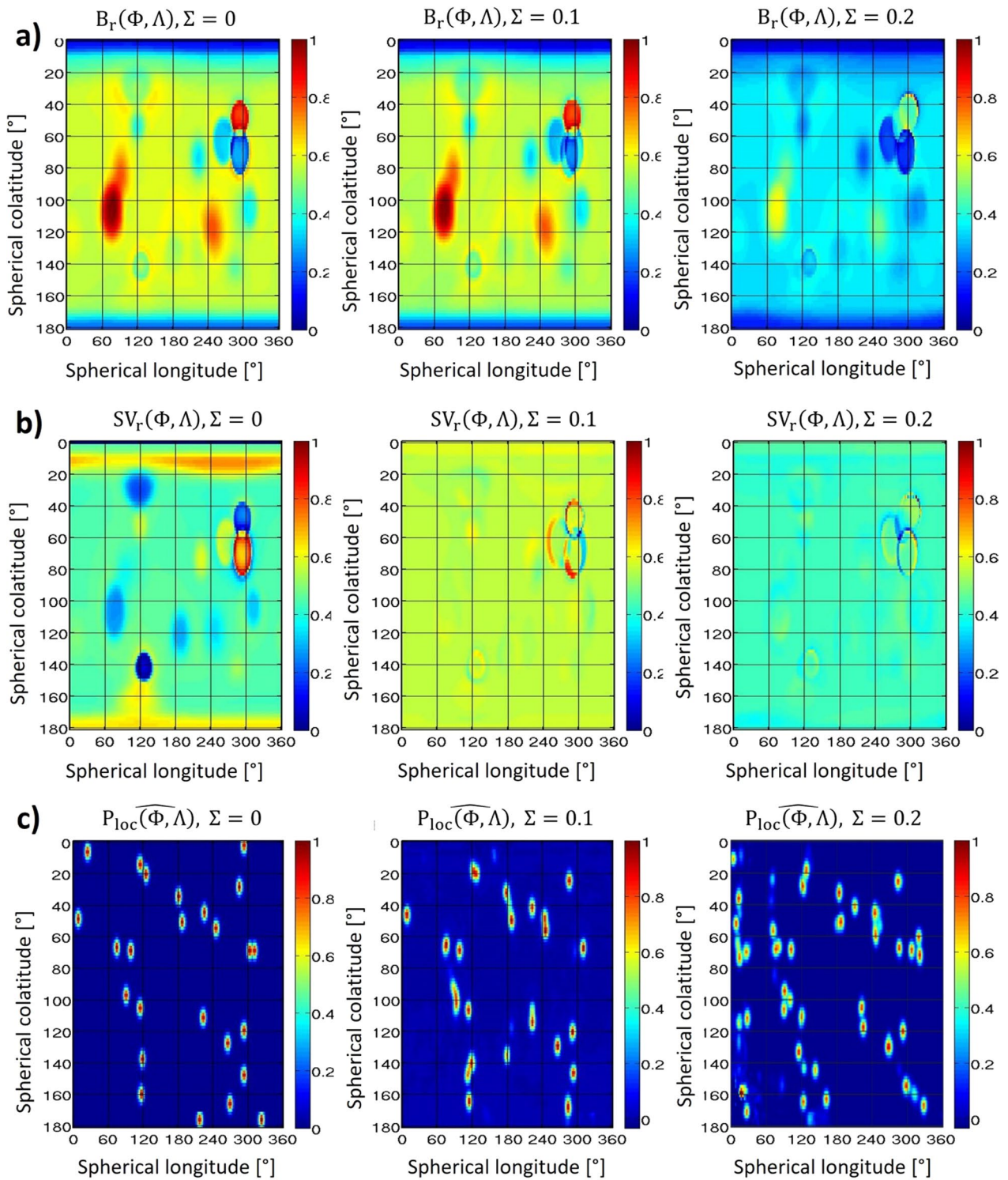


Fig. 30 Response of the inference to noise introduced in the input magnetic induction (a) and SV (b) data. Row c) shows source positions inferred by the network as the noise level Σ increases from left to right

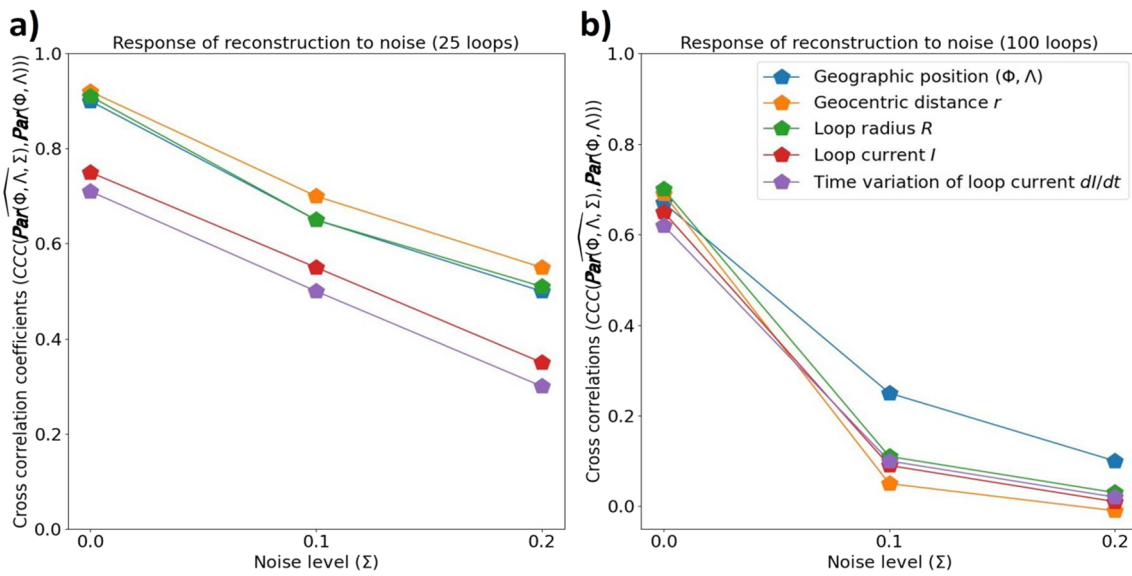


Fig. 31 Cross-correlations between true and inferred source parameter maps plotted against an increasing noise level in a model containing 25 (a) and 100 (b) loops

arise. Most notably, Alldredge mentions, that (at the time) he had no real option for inferring the potential number of current loops and determining an initial set

of loop parameters, other than resorting to intuition, and heuristic considerations.

He tried to address the issue of correct initial placement by assigning initial geographic loop positions via

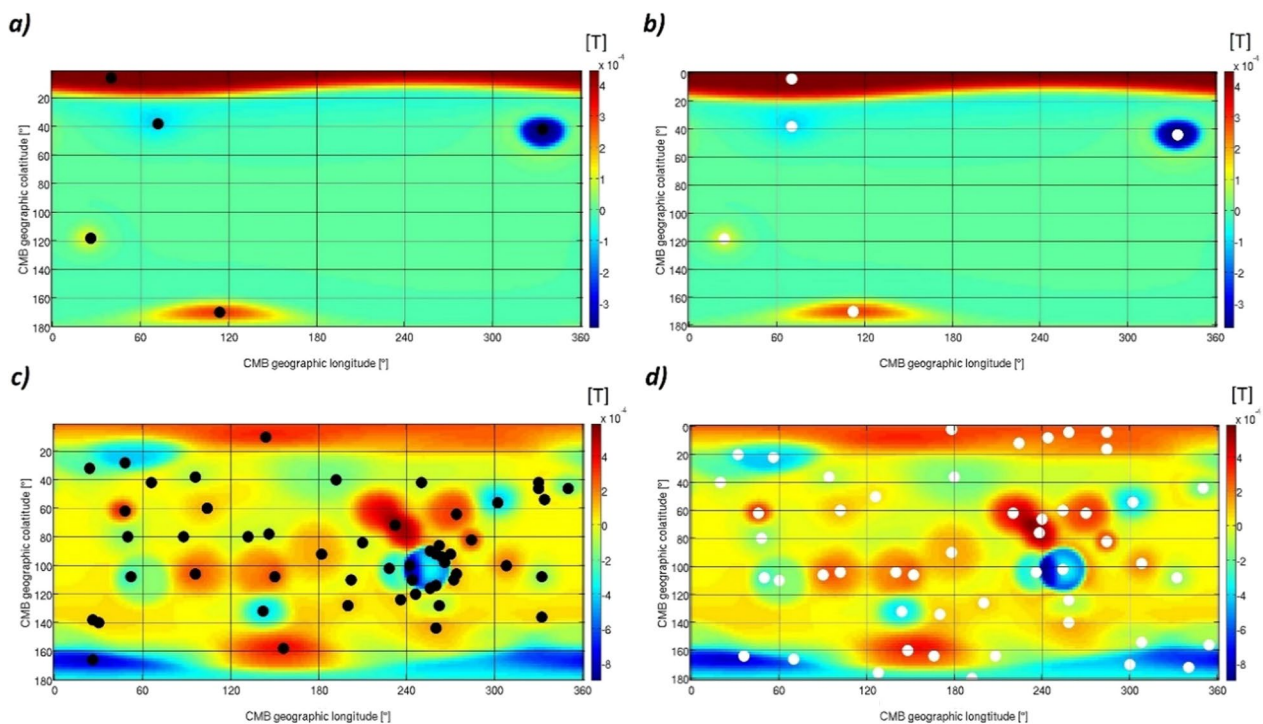


Fig. 32 Initial loop positions assigned using peaks of radial magnetic field values in the model containing the five loops used in the stability analysis (a) and the corresponding correct (actual) positions of the loops (b). Initial loop positions assigned similarly for a model containing 50 loops (c) and the correct loop positions (d)

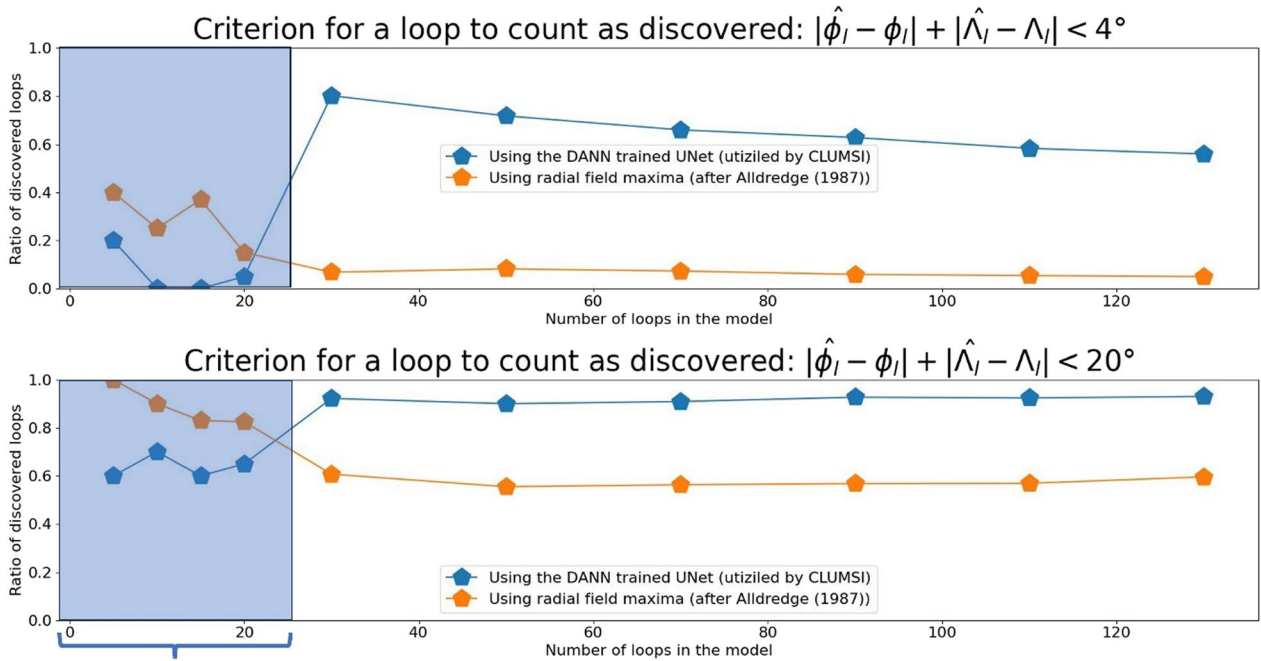


Fig. 33 Average rate of correctly identified current loops as a function of the number of loops in the test examples

identifying ‘peaks’ and ‘valleys’ on maps of the radial geomagnetic field. One can think of a minimum gradient-based procedure, such as the one used in Sect. “Refining the ML-based inversion” to pinpoint source locations on the maps inferred by the network, applied directly on the radial field data to provide an intuitive guess on initial loop parameters. However, as we could already observe in Fig. 6 (Sect. “Approximating the total field assuming linear time variation in the source currents”), the more loops are admitted to a given model, the more complex the corresponding CMB radial fields become. Figures 32 and 33 show that accordingly, it becomes progressively more difficult to assess an initial geographic distribution of the loops correctly.

Figure 33 shows a quantitative comparison of this procedure against our neural net inference on the same data used in part to test our method in Sect. 4.1 (see Figs. 10 and 11). For each sample, the number of correctly identified or “discovered” loops was determined using the criterion:

$$L_{disc_l} = \begin{cases} 1 : \left| \hat{\phi}_l - \phi_l \right| + \left| \hat{\Lambda}_l - \Lambda_l \right| < 4^\circ \\ 0 : \left| \hat{\phi}_l - \phi_l \right| + \left| \hat{\Lambda}_l - \Lambda_l \right| \geq 4^\circ \end{cases} \quad (78)$$

(a distance of 2 gridpoints on the maps) in accordance with condition (1) in Sect. “Approximating the total field assuming linear time variation in the source currents”.

To understand the results of this comparison, we need to consider two things. Firstly, loop models with very few sources have not been shown at all to the network during training. Secondly, assigning loop positions simply via radial field maxima may work reasonably well when the radial fields are generated by a few isolated loops. Thus, in these cases (marked by the blue area on the graphs), the network is much less accurate, while after more and more loops are admitted to the models, we see that the Allredge’s intuition for initial placement becomes

Table 9 Relative change in RPE over the course of $i = 5$ iterations as the ratio of initial RPE

Number of loops in the model	Relative deviations of the geographic placement (ϕ_k, Λ_k) of the loops	RPE ratio ($\Delta RPE/RPE_0$)
5 ^(a)	5e-2	-1.15e-1
5 ^(b)	1e-1	2.08e-2
10 ^(a)	5e-2	-3.32e-1
10 ^(b)	1e-1	-1.2e-1
20 ^(a)	5e-2	2.55e-1
20 ^(b)	1e-1	1.97e-2

(a), (b) refer to the corresponding deviations in geographic placements featured also in Fig. 35 for the same estimations

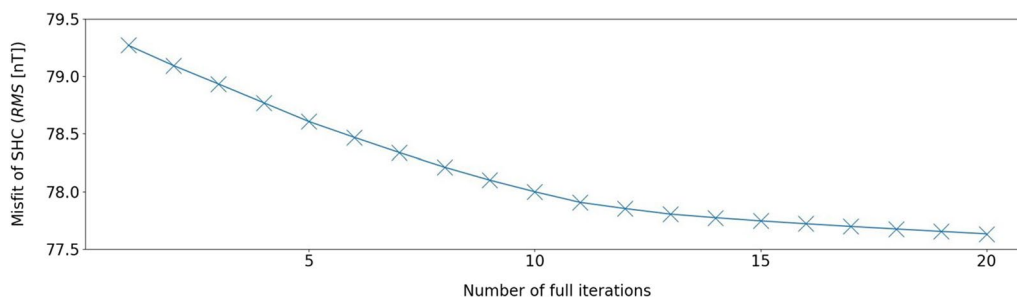


Fig. 34 Graph showing the decrease of RMS error when using Allredge’s method for model 5^(b). The parameters can easily diverge from the correct ones even if the RMS misfit decreases

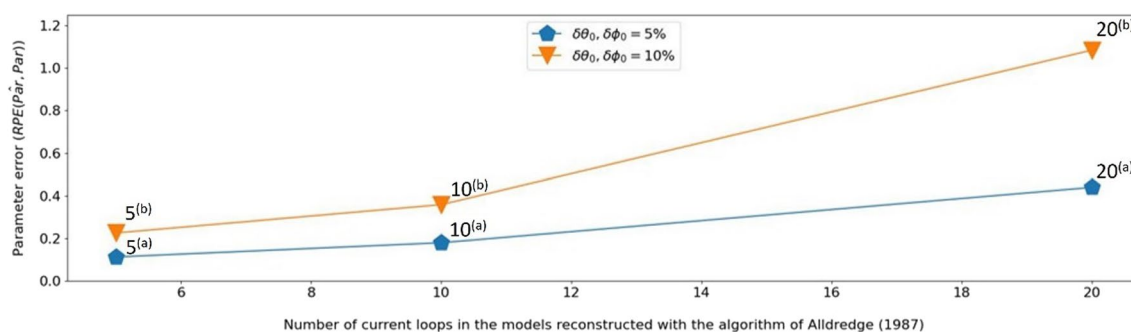


Fig. 35 RPE measures obtained after 5 full iterations assuming different relative errors for the initial placement of current loops using Allredge’s method as a function of the number of loops used in the test examples shown in Table 9

very inaccurate compared to the neural network-based solution.

Moreover, a stability analysis demonstrating the sensitivity of Allredge’s estimation procedure with respect to these initially assigned loop parameters ϕ_l, Λ_l was carried out. During this test, a maximum SH degree of 15 was used in the forward computations. Different deviations from the correct parameters in relation with their absolute values were inspected. A total number of $i = 5$ iterations, carrying out the linear and non-linear estimation steps sequentially as described by Allredge, was performed for each case. The RPE error measure was applied in the same fashion as defined in Eq. (29) in Sect. “Loss functions and quality metrics”. Its respective changes during the total number of iterations ($\Delta RPE = RPE_0 - RPE_i$) are shown in Table 9 for each case.

Table 9 shows that the method tends to stick in local minima when the initial geographic placement of the loops is incorrect. Even comparatively minor misplacements of a relatively small set of five loops resulted in a divergence (positive RPE ratios) of the estimated loop parameters from correct ones even though the corresponding RMS misfit values decreased considerably (see Fig. 34).

Erroneous estimations arise in particular, because a relatively small initial deviation in the initial parameters can result in very high RPE values mostly due to poor estimations for K_l (see Fig. 35).

This again may not be a major issue for recovering a small set of well isolated loops, where a good initial guess for the placement of the loops is possible. However, in a problem setting as the one established in our study, where a potentially very large set of current loops with a wide range of loop parameters can account for the complex pattern of the observed core field, it renders the application of such estimation procedures infeasible. Comparing these results with the RPE and CCC values obtained using our method shown in Figs. 10a and 11, and considering the practical arguments set out in Sect. “Introduction” make the above conclusion even more clear.

Abbreviations

- CLUMSI Current loop-based model segmentation inference
- CMB Core–mantle boundary
- GA Genetic algorithm
- GMD Geomagnetic dipole moment

ML	Machine learning
MHD	Magneto hydrodynamics
SH	Spherical harmonics
SHC	Spherical harmonic coefficients
SV	Secular variation

Acknowledgements

We wish to thank Hannah Rodgers (ISTerre) for helping us in efficiently utilizing COVOBS geomagnetic model data and for her suggestions on the manuscript suggestions. We also wish to thank Antal Jakovác (ELTE Institute of Physics and Astronomy), Csaba Szabó (ELTE Institute of Geography and Earth Sciences), László Balázs (Wigner Research Centre for Physics) and Péter Pollner (ELTE Institute of Physics and Astronomy) for their insightful suggestions. Gábor Hatos is gratefully acknowledged for his prompt availability concerning computational infrastructure at HUN-REN EPSS. We say our thanks to László Szarka, the former head of HUN-REN Research Centre for Astronomy and Earth Sciences for his continuous support concerning this research.

Author contributions

LK invented the concept of loop model solutions, computed the solutions, proposed the application of neural networks and the DANN method, generated the training data along with developing the related computations (in MATLAB), performed the full parameter GA estimations (also in MATLAB) and wrote the bulk of the manuscript. AH constructed the neural network framework, developed the corresponding (python) code base, performed its training and the inferences. VW contributed to the Introduction section via summarizing previous attempts at constructing source models of the geomagnetic field. CB mostly to the Introduction section via summarizing core surface flow estimations and inverse geodynamo models and the Discussion section regarding possible future comparisons between snapshots of inferred loop distributions and core surface flow estimations. TR designed the numerical simulations in COMSOL. EP contributed mostly to the reimplementation of Allredge's method and the corresponding non-linear least-squares estimation and the computational framework for the GA estimations along with the methodology for simulating the induction screening around the loops. LCz established the calculations deriving potential loop parameter values and contributed to finalizing the formatting of the manuscript. IB contributed mostly to the methodology for simulating the induction screening around the loops along with writing the bulk of the overview of current loop representations in the literature. IL invented the computational framework simulating the induction screening around the loops and established the overall algorithmic logic of that framework. All co-authors read and approved the final manuscript.

Funding

Open access funding provided by HUN-REN Institute of Earth Physics and Space Science. This work was supported by the "Machine Learning Methods in Earth Physics" project implemented at the "Institute of Earth Physics and Space Science" under grant SA 46/ 2021 of the Hungarian Research Network (HUN-REN).

Availability of data and materials

Codes are available at <https://github.com/lukacskuslits/kuslits-et-al-current-loops>. Test data, network weights and figures are available at <https://github.com/lukacskuslits/kuslits-et-al-current-loops/blob/main/current-loops-data.zip>. GUFM-1 data are available from <https://github.com/martinrehfeld/gufm1-webservice> COVOBS data are available from <https://geodyn.univ-grenoble-alpes.fr/>.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

¹HUN-REN Institute of Earth Physics and Space Science, Csatkai E. U. 6-8, 9400 Sopron, Hungary. ²Faculty of Information Technology and Bionics, Pazmany Peter Catholic University, 1088 Budapest, Hungary. ³British Geological Survey, Lyell Centre, Riccarton, Edinburgh EH14 4AP, UK. ⁴Quantectum AG, Churerstrasse 80, 8808 Pfäffikon, Switzerland.

Received: 6 October 2023 Accepted: 4 May 2024

Published online: 22 May 2024

References

- Alken P, Thébault E, Beggan CD et al (2021) Evaluation of candidate models for the 13th generation International Geomagnetic Reference Field. *Earth Planet Space*. <https://doi.org/10.1186/s40623-020-01281-4>
- Allredge LR (1987) Current loops fitted to geomagnetic model spherical harmonic coefficients. *J Geomag Geoelectr* 39:271–296
- Aubert J (2019) Approaching Earth's core conditions in high-resolution geodynamo simulations. *Geophys J Int* 219(Supplement 1):137–151. <https://doi.org/10.1093/gji/ggz232>
- Aubert J (2023) State and evolution of the geodynamo from numerical models reaching the physical conditions of Earth's core. *Geophys J Int* 235(1):468–487. <https://doi.org/10.1093/gji/ggad229>
- Aubert J, Gastine T, Fournier A (2017) Spherical convective dynamos in the rapidly rotating asymptotic regime. *J Fluid Mech* 813:558–593. <https://doi.org/10.1017/jfm.2016.789>
- Baker JE (1985) Adaptive selection methods for genetic algorithms. In *Proceedings of the 1st International Conference on Genetic Algorithms*. L. Erlbaum Associates Inc, pp. 101–111
- Basir S, Senocak I (2022) Physics and equality constrained artificial neural networks: application to forward and inverse problems with multi-fidelity data fusion. *J Comput Phys*. <https://doi.org/10.1016/j.jcp.2022.111301>
- Christensen RU (2011) Geodynamo models: tools for understanding properties of Earth's magnetic field. *Phys Earth Planet Inter* 187(3–4):157–169. <https://doi.org/10.1016/j.pepi.2011.03.012>
- Czirok L, Kuslits L, Bozsó I, Radulian M, Gribovszki K (2022) Cluster analysis for the study of stress patterns in the Vrancea-Zone (SE-Carpathians). *Pure Appl Geophys* 179:3693–3712. <https://doi.org/10.1007/s00024-022-03159-w>
- De Santis A, Quamili E (2010) Equivalent monopole source of the geomagnetic South Atlantic Anomaly. *Pure Appl Geophys* 167:339–347. <https://doi.org/10.1007/s00024-009-0020-5>
- Dong C, Zhang H, Jiao L, Cheng L, Yuen DA, Shi Y (2021) The non-negligible effect of viscosity diffusion on the geodynamo process. *J Geophys Res Solid Earth*. <https://doi.org/10.1029/2020JB021281>
- Dongxiao L, Fan W, Shan L, Tian C, Bi B, Zhang F et al (2019) Ab initio simulations for expanded gold fluid in metal-nonmetal transition regime. *Phys Plasmas* 10(1063/1):5123512
- Dutykh D (2016) How to overcome the Courant-Friedrichs-Lewy condition of explicit discretizations? Numerical methods for diffusion phenomena in building physics. https://doi.org/10.1007/978-3-030-31574-0_5
- Ganin Y, Ustinova E, Ajakan H, Germain P, Larochelle H, Laviolette F et al (2016) Domain-adversarial training of neural networks. *J Mach Learn Res* 17(1):2030–2096
- Gillet N, Jault D, Canet E, Fournier A (2010) Fast torsional waves and strong magnetic field within the Earth's core. *Nature* 465(7294):74–77. <https://doi.org/10.1038/nature09010>
- Gillet N, Schaeffer N, Jault D (2011) Rationale and geophysical evidence for quasi-geostrophic rapid. *Phys Earth Planet Inter* 187(3–4):280–390. <https://doi.org/10.1016/j.pepi.2011.01.005>
- Gillet N, Huder L, Aubert J (2019) A reduced stochastic model of core surface dynamics based on geodynamo simulations. *Geophys J Int* 219(1):522–539. <https://doi.org/10.1093/gji/ggz313>
- Glatzmaier GA, Olson P (2005) Probing the geodynamo. *Sci Am* 292(4):50–57
- Grefenstette JJ (1981) Parallel adaptive algorithms for function optimization. *Tech Rep* 18–81
- Gubbins D (1996) A formalism for the inversion of geomagnetic data for core motions with diffusion. *Phys Earth Planet Inter* 98(3–4):193–206. [https://doi.org/10.1016/S0031-9201\(96\)03187-1](https://doi.org/10.1016/S0031-9201(96)03187-1)
- Gubbins D, Roberts PH (1987) Magnetohydrodynamics of the Earth's core. In *Jacobs JA (Ed) Geomagnetism, Vol. 2*, pp 1–183
- Gwirtz K, Davis T, Morzfeld M, Constable C, Fournier A, Hulot G (2022) Can machine learning reveal precursors of reversals of the geomagnetic axial dipole field? *Geophys J Int* 231(1):520–535. <https://doi.org/10.1093/gji/ggac195>

- Hoerl AE, Kennard RW (1970) Ridge regression: biased estimation for non-orthogonal problems. *Technometrics* 12(1):55–67. <https://doi.org/10.2307/1267351>
- Holme R, Olson P, Schubert G (2015) Large-scale flow in the core. In *Treatise on geophysics*. Amsterdam: Elsevier, Vol. 8, pp 107–130
- Huder L, Gillet N, Thollard F (2019) pygeodyn 1.1.0: a Python package for geomagnetic data assimilation. *Geosci Model Dev* 12:3795–3803. <https://doi.org/10.5194/gmd-12-3795-2019>
- Huyen C (2022) Designing machine learning systems. O'Reilly Media, Inc., pp 83, 348
- Ilyas IF, Chu X (2019) Data cleaning. New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/3310205>
- Jackson A, Jonkers AR, Walker MR (2000) Four centuries of geomagnetic secular variation from historical records. *Phil Trans R Soc Lond* 358(1678):957–990. <https://doi.org/10.1098/rsta.2000.0569>
- Jackson JD (1998) Vector potential and magnetic induction for a circular current loop. In *Classical electrodynamics* (pp. 181–184). ISBN: 978-0-471-30932-1
- Jacobs C, van der Holst B, Poedts S (2007) Comparison between 2.5D and 3D simulations of coronal mass ejections. *Astron Astrophys* 470:359–365. <https://doi.org/10.1051/0004-6361:20077305>
- Jarolim R, Thalmann JK, Veronig AM, Podladchikova T (2023) Probing the solar coronal magnetic field with physics-informed neural networks. *Nat Astron* 7:1171–1179. <https://doi.org/10.1038/s41550-023-02030-9>
- Korte M, Constable CG (2005) The geomagnetic dipole moment over the last 7000. *Earth Planet Sci Lett* 236(1–2):348–358. <https://doi.org/10.1016/j.epsl.2004.12.031>
- Kuslits L, Lempereger I, Horváth A, Koronczay D, Wesztergom V (2020) Recent progress in identification of the geomagnetic signature of 3D outer core flows. *Acta Geod Geophys* 55:347–370. <https://doi.org/10.1007/s40328-020-00307-3>
- Ladynin AV (2014) Dipole sources of the main geomagnetic field. *Russ Geol Geophys* 55(4):495–507. <https://doi.org/10.1016/j.rgg.2014.03.007>
- Ling SJ, Moebes W, Sanny J (2016) 12.4 Magnetic field of a current loop. Retrieved from University Physics Volume 2: <https://pressbooks.online.ucf.edu/osuniversityphysics2/chapter/magnetic-field-of-a-current-loop/>
- Loftin S, Fite SJ, Bishop LV, Kotsiaros S (2019) Machine learning vs conventional analysis techniques for the earth's magnetic field study. *SMU Data Sci Rev* 2(1)
- Mayhew MA, Estes RH (1983) Equivalent source modeling of the core magnetic field using Magsat data. *J Geomagn Geoelectr* 35(4):119–130. <https://doi.org/10.5636/jgg.35.119>
- Metman MC, Livermore PW, Mound JE, Beggan CD (2019) Modelling decadal secular variation with only magnetic diffusion. *Geophys J Int* 219(Supplement 1):S58–S82. <https://doi.org/10.1093/gji/ggz089>
- Miralles S, Bonnefoy N, Bourgoin M, Odier P, Pinton J-F, Plihon N et al (2013) Dynamo threshold detection in the von Kármán sodium experiment. *Phys Rev E* 88(1):013002. <https://doi.org/10.1103/PhysRevE.88.013002>
- Miyagoshi T, Kageyama A, Sato T (2008) Formation of current coils in geodynamo simulations. *Nature* 454:1106–1109. <https://doi.org/10.1038/nature07227>
- Miyagoshi T, Kageyama A, Sato T (2011) Formation of sheet plumes, current coils, and helical magnetic fields in a spherical magnetohydrodynamic dynamo. *Phys Plasmas* 10(1063/1):3603822
- Monchoux R, Berhanu M, Aumaitre S, Chiffaudel A, Daviaud F et al (2010) The von Karman Sodium experiment: turbulent dynamical dynamos. *Phys Fluids* 21(3):035108. <https://doi.org/10.1063/1.3085724>
- Mühlenbein H, Schlierkamp-Voosen D (1993) Predictive models for the breeder genetic algorithm I. Continuous parameter. *Optim Evol Comput* 1(1):25–49. <https://doi.org/10.1162/evco.1993.1.1.25>
- Müller U, Stieglitz R, Busse FH, Tilgner A (2008) The Karlsruhe two-scale dynamo experiment. *C R Phys* 9(7):729–740. <https://doi.org/10.1016/j.crhy.2008.07.005>
- MULTIFRONTAL MASSIVELY PARALLEL SOLVER (2022) Retrieved from <https://mumps-solver.org/index.php?page=doc>.
- Multiphysics, COMSOL (1998) Introduction to comsol multiphysics®. COMSOL Multiphysics, Burlington, MA, Accessed Feb, 9(2018), 32
- Ohta K, Kuwayama Y, Hirose K, Shimizu K, Ohishi Y (2016) Experimental determination of the electrical resistivity of iron at earth's core conditions. *Nature* 534(7605):95–98. <https://doi.org/10.1038/nature17957>
- Peddie NW (1979) Current loop models of the Earth's magnetic field. *J Geophys Res* 84(B9):4517–4523. <https://doi.org/10.1029/JB084iB09p04517>
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O et al (2011) Scikit-learn: machine learning in Python. *J Mach Learn Res* 12:2825–2830
- Picek S, Jakobovic D, Golub M (2013) On the recombination operator in the real-coded genetic algorithms. 2013 IEEE Congress on Evolutionary Computation. pp. 3103–3110. <https://doi.org/10.1109/CEC.2013.6557948>
- Popescu SG, Sharp DJ, Cole JH, Kamnitsas K, Glocker B (2021) Distributional Gaussian process layers for outlier detection in image segmentation. *Information Processing in Medical Imaging: 27th International Conference, IPMI 2021* (pp. 415–427). Virtual Event: Springer International Publishing
- Raissi M, Perdikaris P, Karniadakis GE (2019) Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys* 378:686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- Rehfeld M (2019) gufm1-webservice. Retrieved from GitHub: <https://github.com/martinrehfeld/gufm1-webservice>
- Rong ZJ, Wei Y, Lucy Klinger M, Yamauchi WYX, Kong DL, Cui J, Shen C, Yang YY, Zhu RX, Zhong J, Chai LH (2021) A new technique to diagnose the geomagnetic field based on a single circular current loop model. *J Geophys Res Solid Earth*. <https://doi.org/10.1029/2021JB022778>
- Sanchez S, Wicht J, Baerenzung J, Holschneider M, Aubert J, Fournier A (2018) Probing the Earth's core dynamics through geomagnetic observations and dynamo simulations. EGU General Assembly Conference Abstracts, (p. 14663). Retrieved from <https://ui.adsabs.harvard.edu/abs/2018EGUGA.2014663S>
- Schaeffer N, Jault D, Nataf H, Fournier A (2017) Turbulent geodynamo simulations: a leap towards Earth's core. *Geophys J Int* 211(1):1–29. <https://doi.org/10.1093/gji/ggx265>
- Sheyko A, Finlay CC, Jackson A (2016) Magnetic reversals from planetary dynamo waves. *Nature* 539:551–554. <https://doi.org/10.1038/nature19842>
- Simpson JC, Lane JE, Immer CD, Youngquist RC (2001) Simple analytic expressions for the magnetic field of a circular current loop. No. NASA/TM-2013-217919
- Su S, Cébron D, Nataf H-C, Cardin P, Vidal J, Solazzo M, Do Y (2020) Acoustic spectra of a gas-filled rotating spheroid. *Eur J Mech B-Fluids* 84:302–310. <https://doi.org/10.1016/j.euromechflu.2020.03.003>
- Talagrand O (1997) Assimilation of observations: an introduction. *J Meteorol Soc Jpn Ser II* 75(1B):191–209. https://doi.org/10.2151/jmsj1965.75.1B_191
- Weaver JT (1994) 2.4 Homogeneous half-space. In *Mathematical methods for geo-electromagnetic induction* (p. 35). Research Studies Press
- Whaler KA (1986) Geomagnetic evidence for fluid upwelling at the core-mantle boundary. *Geophys J Int* 86(2):563–588. <https://doi.org/10.1111/j.1365-246X.1986.tb03844.x>
- Whaler KA, Beggan CD (2015) Derivation and use of core surface flows for forecasting secular variation. *J Geophys Res Solid Earth* 120(3):1400–1414. <https://doi.org/10.1002/2014JB011697>
- Wieczorek MA, Meschede M (2018) SHTools: tools for working with spherical harmonics. *Geochem Geophys Geosyst* 19(8):2574–2592. <https://doi.org/10.1029/2018GC00752>
- Yang WY, Cao W, Chung T-S, Morris J (2005) Applied numerical methods using MATLAB®. Wiley. <https://doi.org/10.1002/0471705195>.
- Zhou Z, Rahman SM, Tajbakhsh N, Liang J (2018) U-net++: A nested u-net architecture for medical image segmentation. Deep learning in medical image analysis and multimodal learning for clinical decision support: 4th International Workshop. 4, pp. 3–11. Granada: Springer International Publishing. https://doi.org/10.1007/978-3-030-00889-5_1

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.