Machine Learning for the geodynamo inverse problem

Physics-informed Neural Network

Romain Claveau

May 30, 2025

Table of Contents

1 Context

The geodynamo is a physical process behind the Earth's sustained magnetic field, stemming from complex fluid motions of the electrically conducting liquid metal within the outer core. Because these processes occur at extreme depth and conditions, which are entirely out-of-reach of direct observations or experimental replicas, our understanding of the geodynamo relies heavily on numerical simulations.

However, the Earth's physical parameters are also out-of-reach of numerical simulations as its outer core operates at low viscosity (as measured by $\mathrm{Ek} \simeq 10^{-15}$ and $\mathrm{Pm} \simeq 10^{-6}$). Also, the Earth's dynamo features a small ratio of the kinetic to magnetic energy (as measured by $\mathrm{Al}^2 \simeq 10^{-4}$). So, reproducing Earth-like conditions requires to reach an asymptotic regime, which afterwards could be extrapolated to the Earth.

We recall the definitions of the Ekman number (Ek), the magnetic Prandlt number (Pm) and the Alfvén number (Al):

$$\mathrm{Ek} = \frac{\tau_\Omega}{\tau_\nu} \;\; ; \;\; \mathrm{Pm} = \frac{\tau_\eta}{\tau_\nu} \;\; ; \;\; \mathrm{Al} = \frac{\tau_A}{\tau_U} \tag{1.1}$$

where τ_Ω the inverse rotation rate, τ_ν the viscous diffusion time, τ_η the magnetic diffusion time, τ_A the Alfvén time and τ_U the convective overturn time.

As a result, in addition to numerical simulations of the geodynamo, we rely on the induction equation relating the outer core motions and the magnetic field, to infers the flow through magnetic observations by solving the geodynamo inverse problem.

2 The geodynamo inverse problem

The induction equation, relating the core flow to the magnetic field, is obtained from Ohm's law (accounting for the Lorentz force)

$$\boldsymbol{J} = \sigma(\boldsymbol{E} + \boldsymbol{u} \times \boldsymbol{B}) \tag{2.1}$$

where $\boldsymbol{J}$ is the electric current field, $\boldsymbol{E}$ the electrical field, $\boldsymbol{u}$ the velocity field and $\boldsymbol{B}$ the magnetic field. Because the Earth's mantle is electrically insulating (at least has a very low electrical conductivity), $\boldsymbol{J}$ vanishes. Then, taking the curl, and using the Faraday's law, one gets the ideal magneto-hydrodynamics equation

$$\frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla} \times (\boldsymbol{u} \times \boldsymbol{B}) \tag{2.2}$$

where $\boldsymbol{\nabla}\times$ is the curl operator. Also, because we are considerating the mantle as insulating, the magnetic field is divergence-free and curl-free. As a result, it derives from a potential $\boldsymbol{B} = -\boldsymbol{\nabla} V$, with V the magnetic potential, and its evolution above the outer core surface is entirely prescribed by its radial component at the core-mantle boundary (CMB).

At the CMB, the outer core motions are constrained by the mantle as it cannot flow radially. As a result, only its tangential components are non-zero. Finally, because the radial magnetic field is the only one driven by these tangential motions, one gets the radial induction equation at the CMB, reading as

$$\frac{\partial B_r}{\partial t} = -\boldsymbol{\nabla}_H \cdot (\boldsymbol{u}_H B_r) \tag{2.3}$$

where $\boldsymbol{\nabla}_H \cdot$ is the horizontal (tangential) divergence operator, B_r the radial magnetic field and $\boldsymbol{u}_H$ the tangential flow.

Thus, the *geodynamo inverse problem* is precisely inferring the tangential outer core flow $\boldsymbol{u}_H$ from the magnetic observations B_r.

3 Tangential flow field

As the outer core flow is incompressible, it admits a unique toroidal–poloidal decomposition:

$$\boldsymbol{u} = \boldsymbol{\nabla} \times (\boldsymbol{r}\,\mathcal{T}) + \boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times (\boldsymbol{r}\,\mathcal{S})) \tag{3.1}$$

with $\mathcal{T}$ and $\mathcal{S}$ the toroidal and poloidal scalar fields, respectively. The expression is further simplied as we are considering the tangential flow, and reads

$$\boldsymbol{u}_H = -\boldsymbol{r} \times \boldsymbol{\nabla}_H \mathcal{T} + \boldsymbol{\nabla}_H (r\,\mathcal{S}) \tag{3.2}$$

Using this expression for the tangential flow field $\boldsymbol{u}_H$ ensures the incompressibility condition (ie. $\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$).

4 Inverting for the tangential core flow

The *geodynamo inverse problem* is an ill-defined problem as we have to retrieve the two tangential components of the core flow, u_θ and u_ϕ, using one equation. In this context, we have to provide additional information to ensure the uniqueness of the solution. Although this is usually done using Bayesian-like optimization, the use of *physics-informed neural network* (PINN) could be a way forward, allowing (hopefully) to resolve small length-scales that are computationally prohibitive otherwise.

4.1 Framework

We are solving the radial induction equation at the CMB

$$\frac{\partial B_r}{\partial t} = -\boldsymbol{\nabla}_H \cdot (\boldsymbol{u}_H B_r) \tag{4.1}$$

where $\partial_t B_r$ and B_r are extracted from already-existing magnetic models, such as `COV-OBS.x2` or `CHAOS-7`.

The incompressibility condition on the tangential core flow $\boldsymbol{u}_H$ is enforced by considering a toroidal–poloidal decomposition.

$$\boldsymbol{u}_H = -\boldsymbol{r} \times \boldsymbol{\nabla}_H \mathcal{T} + \boldsymbol{\nabla}_H (r\, \mathcal{S}) \tag{4.2}$$

The flow can further be constrained, by having it to be quasi-geostrophic (ie. if the force balance is dominated by Coriolis forces, then the flow will align itself along the rotation axis, creating a columnar-like flow). This condition reads

$$\boldsymbol{\nabla}_H \cdot (\boldsymbol{u}_H \cos(\theta)) = 0 \tag{4.3}$$

which can further be expanded by distributing the horizontal divergence over the product, as

$$\boldsymbol{\nabla}_H \cdot \boldsymbol{u}_H - u_\theta \frac{\tan(\theta)}{r} = 0 \tag{4.4}$$

Then, we are defining two loss functions, L_1 and L_2 to quantify the misfits to data (L_1) and the quasi-geostrophy (L_2) as

$$\left\{\begin{array}{lcl} L_1 &=& \frac{1}{N}\sum_{\text{grid}} \left\| \frac{\partial B_r}{\partial t} + \boldsymbol{\nabla}_H \cdot (\boldsymbol{u}_H B_r) \right\|^2 \\ L_2 &=& \frac{1}{N}\sum_{\text{grid}} \left\| \boldsymbol{\nabla}_H \cdot \boldsymbol{u}_H - u_\theta \frac{\tan(\theta)}{r} \right\|^2 \end{array}\right. \tag{4.5}$$

The *best* model is that minimizing the total loss function

$$\arg\min_{\lambda} L_1(\theta,\phi) + \lambda L_2(\theta,\phi) \tag{4.6}$$

4.2 The equation and fields in spherical coordinates

First, we start by expressing the toroidal and poloidal parts of the tangential flow field. We define $\boldsymbol{u}_T := -\boldsymbol{r} \times \boldsymbol{\nabla}_H \mathcal{T}$ and $\boldsymbol{u}_S := \boldsymbol{\nabla}_H (r\mathcal{S})$. Consequently, these two fields read

$$\left\{\begin{array}{lcl} \boldsymbol{u}_T &=& \left(\frac{1}{\sin(\theta)} \frac{\partial \mathcal{T}}{\partial \phi}, -\frac{\partial \mathcal{T}}{\partial \theta} \right) \\ \boldsymbol{u}_S &=& \left(\frac{\partial \mathcal{S}}{\partial \theta}, \frac{1}{\sin(\theta)} \frac{\partial \mathcal{S}}{\partial \phi} \right) \end{array}\right. \tag{4.7}$$

Recalling the induction equation, we need to express the horizontal divergence of the tangential flow in spherical coordinates.

$$\boldsymbol{\nabla}_H \cdot \boldsymbol{u}_H = \frac{1}{r\sin(\theta)} \frac{\partial}{\partial \theta}(u_\theta \sin(\theta)) + \frac{1}{r\sin(\theta)} \frac{\partial u_\phi}{\partial \phi} \tag{4.8}$$

Finally, the induction equation in spherical coordinates reads

$$\begin{array}{rcl} \frac{\partial B_r}{\partial t} &=& -\frac{1}{r\sin(\theta)} \left[\frac{\partial}{\partial \theta}(u_\theta \sin(\theta)) + \frac{\partial u_\phi}{\partial \phi} \right] B_r \\ && -\frac{1}{r\sin(\theta)} \left[u_\theta \sin(\theta) \frac{\partial B_r}{\partial \theta} + u_\phi \frac{\partial B_r}{\partial \phi} \right] \end{array} \tag{4.9}$$

with the quasi-geostrophic constraint

$$\frac{1}{\sin(\theta)}\left[\frac{\partial}{\partial\theta}(u_\theta \sin(\theta))+\frac{\partial u_\phi}{\partial\phi}\right]-u_\theta \tan(\theta)=0 \tag{4.10}$$

4.3 Strategy

Directly inverting for the core flow on the whole grid would be numerically challenging, as we would have to use a large neural network to model small length-scales.

An idea, as done by Shakespeare-Rees et al. (2025), is to generate a patchwork of smaller rectangles. The core flow is then inverted on all the rectangles. Overlaps are added for continuity purpose around all rectangles.

A first optimization is carried out using the Adam algorithm. Then, fine-tuning may be done using the L-BFGS algorithm. Finally, this procedure could be carried out several time, with different initial conditions and averaged.

Regarding the neural network itself, as we are inverting the core flow on a small portion of space at a time, we do not have to rely on large or deep network. A good starting point might be:

i. Number of nodes: 64
ii. Number of hidden layers: 3

In order to get a finer resolution, two solutions are possible:

i. Reduce the size of each rectangle of the patchwork
ii. Increase the size of the neural network

In the first case, it allows *de facto* a better resolution and also to lower the size of the neural network as its associated data should be smoother. In the second case, it might allow a better description of small length-scales but the neural network may fit noise instead.

Appendix A Del in spherical coordinates

Recalling the usual (but useful) formulae of the del operator in spherical coordinates, applied either on a scalar field $F(r,\theta,\phi)$ or a vector field $\boldsymbol{F}(r,\theta,\phi)$.

A.1 Horizontal gradient

$$\boldsymbol{\nabla}_H F = \left(\frac{1}{r}\frac{\partial F}{\partial\theta}, \frac{1}{r\sin(\theta)}\frac{\partial F}{\partial\phi}\right) \tag{A.1}$$

A.2 Horizontal divergence

$$\boldsymbol{\nabla}_H \cdot \boldsymbol{F}_H = \frac{1}{r\sin(\theta)}\frac{\partial}{\partial\theta}(F_\theta\sin(\theta)) + \frac{1}{r\sin(\theta)}\frac{\partial F_\phi}{\partial\phi} \tag{A.2}$$