

Pre-trained Foundation Models for Geophysical Magnetic Inversion Problems

Valeriia Okhmak⁽¹⁾, Hakan Bagci⁽¹⁾, and George Turkiyyah⁽²⁾

⁽¹⁾ Electrical and Computer Engineering (ECE) Program

Division of Computer, Electrical, and Mathematical Science and Engineering (CEMSE)

King Abdullah University of Science and Technology (KAUST)

Thuwal 23955-6900, Saudi Arabia

⁽²⁾ Applied Mathematics and Computational Science (AMCS) Program

Division of Computer, Electrical, and Mathematical Science and Engineering (CEMSE)

King Abdullah University of Science and Technology (KAUST)

Thuwal 23955-6900, Saudi Arabia

Three-dimensional (3D) geophysical imaging, posed as an inverse problem, has been a critical challenge in subsurface reconstruction using surface and geophysical data. This study introduces a novel artificial intelligence (AI)-driven method for inverting magnetic data as a way to address this challenge. More specifically, this method estimates magnetic susceptibility, a key parameter for characterizing subsurface Earth structures, using real total magnetic intensity data. A pre-trained, diffusion-based, foundation model acts as a regularizer, enhancing the reliability and robustness of geological image reconstructions.

The inverse problem can be cast as the following minimization problem,

$$\mathcal{L}(\mathbf{x}; \theta) = \frac{1}{2} \|\mathcal{F}(\mathbf{x}) - \mathbf{y}\|^2 + \lambda R_\theta(\mathbf{x}),$$

where $\mathcal{F}(\mathbf{x})$ represents the forward model that maps the subsurface properties $\mathbf{x}$ to the observable domain, $\mathbf{y}$ denotes the observed magnetic data, $R_\theta(\mathbf{x})$ is the regularization term implemented by the foundation model, parameterized by θ , and λ is a regularization parameter to balance the impact of the regularization term.

Traditional inversion methods often use regularizers with simple analytical expressions. An example of such traditional methods is the magnetic inversion with the classical L1–L2 norm regularization (Utsugi, M., *Earth Planets Space*, 71:73 2019). While these methods work well for problems with simple subsurface structures, they are not powerful enough for inverting complex substructures of practical interest. A data-informed prior in the form of a pre-trained generative model is a step toward more general magnetic inverse solver that scales effectively to complex scenarios.

To train the generative foundation model and to get around the scarcity of publicly available 3D geological images, database of geologically reasonable models are generated synthetically by a kinematic process simulator that converts a sequence of deformation events into a 3D geological model (Jessel M. et al., *Earth Syst. Sci. Data*, 14, 381–392, 2022). The neural backbone of the model is defined as a time-conditional U-Net.

The foundation model integrates a generative model based on stochastic interpolation. This generative model creates image instances by solving the stochastic differential equation (SDE), utilizing the learned drift coefficient $b(t, x_t)$ from the Fokker-Planck equations. These equations characterize the evolution of the stochastic process x_t , which transitions between two arbitrary distributions $p_0(x_0)$ and $p_1(x_1)$, where $p_1(x_1)$ is a target distribution. The distribution of the process x_t at any time t is denoted by $p(t, x_t)$ (Albergo, M. S., Boffi, N. M., and Vanden-Eijnden, E., arXiv preprint arXiv:2303.08797, 2023).