Bias-Corrected Eddy-Current Simulation Using a Recurrent-Neural-Net / Finite-Element Hybrid Model

DOI
https://doi.org/10.18690/um.feri.4.2025.5

ISBN
978-961-286-986-1

Moritz von Tresckow, Herbert De Gersem, Dimitrios Loukrezis

Technical University of Darmstadt, Institute for Accelerator Science and Electromagnetic Fields (TEMF), Darmstadt, Germany
moritz.von_tresckow@tu-darmstadt.de, degersem@temf.tu-darmstadt.de, loukrezis@temf.tu-darmstadt.de

This work combines recurrent neural networks (RNNs) with the finite element (FE) method into a hybrid model to correct time-dependent discrepancies in low-fidelity engineering simulations. The hybrid model is trained on sparse data from high- and low-fidelity simulations, employing techniques to prevent overfitting and balance accuracy with neural network generalization. It is successfully applied to an eddy-current simulation of a quadrupole magnet, demonstrating its accuracy in adjusting low-fidelity models. The results confirm the potential of this hybrid modeling approach for model-based predictions in dynamic multi-fidelity modeling contexts.

Keywords:
multi-fidelity modeling,
bias correction,
recurrent neural netwokrs,
finite element method,
magnet

University of Maribor Press

I Introduction

In the context of multi-fidelity modeling, low-fidelity models inevitably exhibit discrepancies with respect to high-fidelity models, due to systematic errors that arise from modeling assumptions, low mesh resolution, or imperfect physical knowledge. Quantifying these discrepancies is important to assess whether a low-fidelity model is sufficiently accurate, especially, when simulating dynamical systems.

The relationship between a high- and low-fidelity system state at time $t \in [0, T]$, $\boldsymbol{a}_t^{\text{hifi}}$ and $\boldsymbol{a}_t^{\text{lofi}}$, respectively, is given as

$$\boldsymbol{a}_t^{\text{hifi}} = \boldsymbol{a}_t^{\text{lofi}} + \delta_t, \tag{1}$$

where $\delta_t : \mathbb{R}^d \to \mathbb{R}^d$ is a discrepancy function capturing systematic errors [1]. In practical scenarios however, the trajectory data $\boldsymbol{A}_{\text{hifi}} := \{\boldsymbol{a}_{t_k}^{\text{hifi}}\}_{t_k \in T_{\text{hifi}}}$ is only known at specific, finite, time instances $T_{\text{hifi}} := \{t_0, \dots, t_{N_T}\}$, where $N_T \in \mathbb{N}$ denotes the number of time instances. Thus, to account for missing time steps, a parametric model, δ_θ, is necessary to approximate the discrepancy function for systemic error approximation [2].

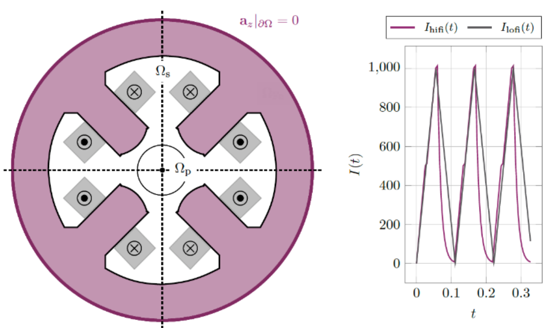

Fig. 1. Left: Schematic of the quadrupole magnet, where $\boldsymbol{\Omega}_{\text{Fe}}$ denotes the domain of the iron yoke, $\boldsymbol{\Omega}_{\text{s}}$ the domain of current excitation and $\boldsymbol{\Omega}_{\text{p}}$ the aperture domain. Right: Current excitation $\mathbf{I}_{\text{hifi}}$ for the high-fidelity model and $\mathbf{I}_{\text{lofi}}$ for the low-fidelity model.

In this work, we propose a framework based on recurrent neural networks (RNNs) and finite element (FE) basis functions, to approximate the model discrepancy $\delta_{\theta*} \approx \delta_t$ in a multi-fidelity, transient, eddy-current simulation of a quadrupole magnet [3].

We employ an upsampling scheme to account for the sparsity of $\boldsymbol{A}_{\text{hifi}}$ and train the model using an upsampled data set. Finally, we use the trained model to derive a correction operator (bias correction), $\boldsymbol{a}^{\text{corr}}$, and improve the performance of the low-fidelity model i.e.

$$\boldsymbol{a}^{\text{corr}} = \boldsymbol{a}_t^{\text{lofi}} + \delta_{\theta^*}, \tag{2}$$

where θ^* is the RNN's trained parameter set.

II FE Model

Choosing the vector potential ansatz $\boldsymbol{b} = \nabla \times \boldsymbol{a}$, with $\boldsymbol{a}$ the magnetic vector potential and $\boldsymbol{b}$ the magnetic flux density, the eddy-current problem can be given as the time-dependent boundary value problem (BVP)

$$\nabla \times (\nu \nabla \times \boldsymbol{a}) + \sigma \frac{\partial \boldsymbol{a}}{\partial t} = \boldsymbol{j}_{\text{s}}\,, \tag{3a}$$

$$\boldsymbol{a}|_{\partial\Omega} = 0\,, \tag{3b}$$

where ν is the reluctivity, σ the conductivity and $\boldsymbol{j}_{\text{s}}$ the source current density.

The geometry of the quadrupole (Fig. 1) is defined on a circular domain Ω, consisting of an iron yoke Ω_{Fe}, coils Ω_{s}, and an aperture Ω_{p}. The domain boundary $\partial\Omega$ consists of the outer boundary of the iron domain $\partial\Omega_{\text{Fe}}$. Due to geometrical considerations, it is sufficient to consider the axial component of the magnetic vector potential, i.e. $\boldsymbol{w}_i = w_i \boldsymbol{e}_z$, with $w_i \in H_0(grad;\Omega)$. The magnetic vector potential is approximated via the ansatz function $\boldsymbol{a} = \sum_{j=1}^{N_{\text{dof}}^{\text{lofi}}} \hat{a}_j \boldsymbol{w}_j$, where the degrees of freedom (dof) $\{\hat{a}_j\}_{j \leq N_{\text{dof}}^{\text{lofi}}}$ lie on the mesh nodes. In matrix-vector notation, the FE formulation reads

$$(\Delta t\, \boldsymbol{A} + \boldsymbol{M})\hat{\boldsymbol{a}}_{t_{k+1}} = \Delta t\, \boldsymbol{b}(t_{k+1}) + \boldsymbol{M}\hat{\boldsymbol{a}}_{t_k}, \tag{4}$$

where $\boldsymbol{A}$ *and* $\boldsymbol{M}$ are the stiffness and mass matrix, respectively, and $\boldsymbol{b}$ is the load vector. For the simulation, we apply the conductivity $\sigma_{Fe} = 1.04 \cdot 10^7\ S/m$ and the reluctivity $\nu_{Fe} = 2 \cdot 10^{-3}\, \nu_0$ in the iron yoke Ω_{Fe}, as well as the conductivity $\sigma = 1\ \text{S/m}$ and the reluctivity $\nu_{\text{Fe}} = \nu_0$ in the aperture Ω_{p} and the current-excitation domain Ω_{s}.

Furthermore, we select a constant time step of $\Delta t = 1 \cdot 10^{-2}$ s and apply $N_T = 327$ time steps. The low-fidelity model $\hat{\boldsymbol{a}}_{t_{k+1}}^{\text{lofi}}$ is parametrized with the current I_{lofi} and mesh resolution $N_{\text{dof}}^{\text{lofi}} = 895$, while the high-fidelity model $\hat{\boldsymbol{a}}_{t_{k+1}}^{\text{hifi}}$ with current I_{hifi} and $N_{\text{dof}}^{\text{lofi}} = 277\,594$.

III Hybrid Model Architecture

To approximate the discrepancy function on a low-fidelity mesh, we assume the functional form

$$\delta_{t_k}(\boldsymbol{r}) = \sum_{i=1}^{N_{\text{dof}}^{\text{lofi}}} \hat{\delta}_{i,t_k} \phi_i(\boldsymbol{r}), \tag{5}$$

where $\hat{\delta}_{i,t_k}$ is the coefficient corresponding to the i-th dof in the low-fidelity mesh and $\phi_i(\boldsymbol{r})$ the associated shape function. We employ an RNN to learn the coefficients of (5) using discrepancy data calculated by

$$D_{\text{d}} := \left\{T\left(\boldsymbol{a}_{t_k}^{\text{hifi}}\right) - \boldsymbol{a}_{t_k}^{\text{lofi}}\right\}_{t_k \in T_{\text{hifi}}}, \tag{6}$$

where T is a linear projection operator, $t_k \in T_{\text{hifi}}$ and the FE shape functions resolve the approximation spatially.

By separating the time-dependent aspects from the spatial elements of the discrepancy function, the RNN focuses solely on temporal coefficient variations, while the FE method handles the spatial discretization of the model domain. This separation allows each method to operate within its area of strength, improving the efficiency of the training process.

IV Localized Data Upsampling

To account for sparsity in the training data, we employ upsampling using localized linear interpolation, i.e.,

$$\bar{\delta}_{j_{t+l}} = \delta_{j_k} + l\left(\frac{\delta_{j_{k+1}} - \delta_{j_k}}{j_{k+1} - j_k}\right), \tag{7}$$

for $l = 1, \dots, j_{k+1} - j_k - 1$, where j_k denotes the time steps for which the high-fidelity data is known. The intermediate artificial system states are coupled with a Gaussian prior to mitigate overfitting in the neural network. This approach, favored for its simplicity and versatility, not only guides NN behavior in sparse data conditions but also uses the Gaussian prior's variance to produce new, artificially bounded states during each training epoch, thus enhancing the model's numerical stability.

V Results

The low-fidelity model uses a much coarse mesh than the high-fidelity model and considers a simplified triangular function for current excitation instead of the true decaying exponential one. The potential distributions obtained with the low-fidelity and bias-corrected models are depicted in Fig. 2a. The integrated discrepancy function over time is shown in Fig. 2b. Relative to the high-fidelity model, the error of the low-fidelity model is $\Delta_{L^2}\boldsymbol{a}^{\text{lofi}} = 39.847\%$, whilst the relative error of the bias-corrected model is $\Delta_{L^2}\boldsymbol{a}^{\text{corr}} = 0.613\%$, which constitutes a major improvement to the low-fidelity model.

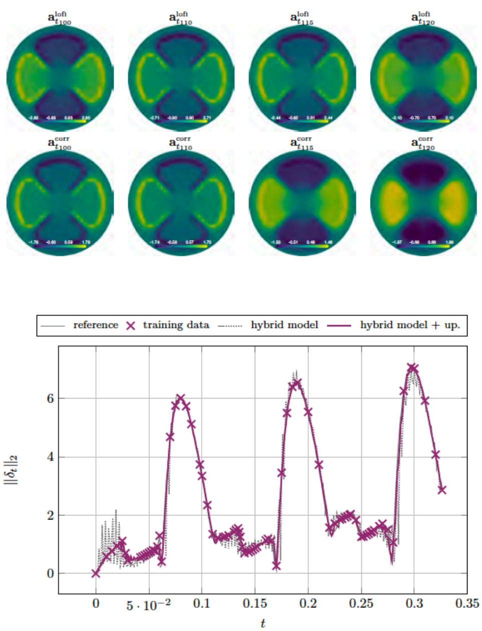

Figure 2: Top: Potential distribution of the low-fidelity and bias-corrected model. Bottom: Integrated discrepancy function over time.

VI Conclusion & Outlook

The hybrid modeling approach yields highly accurate bias-corrected dynamic FE simulations, maintaining error rates below 2% even with irregular behavior and limited data. Future enhancements could come from advanced RNNs, localized Gaussian processes, and tailored, problem-specific loss functions.

Acknowledgment

Moritz von Tresckow acknowledges the support of the German Federal Ministry for Education and Research (BMBF) via the research contract 05K19RDB. Dimitrios Loukrezis and Herbert De Gersem acknowledge the support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project-ID 492661287 – TRR 361. The authors acknowledge the financial support of Deutsches Elektronen-Synchrotron DESY.

References

[1] P. D. Arendt, D. W. Apley, and W. Chen, "Quantification of Model Uncertainty: Calibration, Model Discrepancy, and Identifiability," Journal of Mechanical Design, vol. 134, no. 10, 09 2012, 100908. [Online]. Available: https://doi.org/10.1115/1.4007390.

[2] M. Levine and A. Stuart, "A framework for machine learning of model error in dynamical systems," Communications of the American Mathematical Society, vol. 2, no. 07, pp. 283–344, 2022.

[3] M. von Tresckow, H. De Gersem, and D. Loukrezis, "Error approximation and bias correction in dynamic problems using a recurrent neural network/finite element hybrid model," Applied Mathematical Modelling, 2024.