Efficient full-state estimation via sparse sensors for adaptable GIL Digital Twin application

Junqi Wang[1], Ziqi Zhou[1*], Wensheng Gao[1] Jiangang Ding[2] and Zaixing Peng[2]
[1]Department of Electrical Engineering, Tsinghua University, Beijing, China
wang-jq25@mails.tsinghua.edu.cn, zhouzq23@mails.tsinghua.edu.cn, wsgao@tsinghua.edu.cn
[2]Electric Power Research Institute, China Southern Power Grid, Guangzhou, China
dingjg@csg.cn, pengzx@csg.cn

***Abstract*—Fast and generalizable field reconstruction from sparse sensing is essential for Digital Twin (DT) models, yet maintaining accuracy under varying operational conditions remains a challenge. The challenge is especially pronounced in enclosed equipment like Gas-Insulated Transmission Lines (GIL), where changes in load and environment require models that can generalize effectively. This work employs a parametric Shallow Recurrent Decoder (SHRED) architecture for robust multi-physics field reconstruction in a GIL, using only sparse temperature measurements and key operational parameters. The temperature and velocity fields are reconstructed from noisy measurements across different operational scenarios. The reconstruction uses the temporal dynamics learned by a Gated Recurrent Unit (GRU) and the spatial mapping capabilities of a Shallow Decoder Network (SDN). Results show that the model achieves high reconstruction accuracy, maintaining a Mean Absolute Percentage Error below 0.25% with noise and remains stable under random sensor placement. Its lightweight design and adaptability to varying parameters make it suitable for building robust and efficient real-time DT applications for high-voltage apparatus.**

***Index Terms*—GIL, field reconstruction, parametric model, digital twin, multi-physics simulation**

I. Introduction

Revealing the internal state of high-voltage apparatus from limited external measurements has been a long-standing challenge. With the ongoing digital transformation of modern power systems, Digital Twin (DT) technologies integrated with artificial intelligence (AI) have gained attention as a method to support intelligent operation and maintenance [1], [2]. As key components for high-capacity power transmission, the reliable and efficient operation of Gas-Insulated Transmission Lines (GIL) contributes to overall grid performance [3]. Developing a DT for GIL that can perform full-state estimation in real time with sparse sensor data is a promising direction.

To estimate the multi-physics behavior inside GIL, numerical methods such as the finite element method (FEM) and the finite volume method (FVM) can be employed [4], [5]. Although the simulation method offers high accuracy, it demands significant computational resources and is difficult to apply in real-time monitoring. This has led to the development of Reduced-Order Models (ROMs) and other data-driven surroates that lower computational costs. However, much of the existing research focuses on accelerating forward simulations [6], while the inverse problem, which focuses on reconstructing the full state from measurements, remains a challenge.

Existing techniques for inverse problem have often been static, relying on single-moment measurements and ignoring the rich information in past sensor data. Methods such as Gappy POD also often demand complex optimization algorithms to identify ideal sensor locations for a well-conditioned problem [7]. Additionally, models trained under specific conditions tend to generalize poorly when operating conditions change, thereby limiting their effectiveness in dynamic environments [8].

To overcome these challenges, this work employs a parametric Shallow Recurrent Decoder (SHRED) architecture to efficiently reconstruct a real-time digital twin for GIL [9]–[12]. It is composed of two computationally efficient components: a Gated Recurrent Unit (GRU) for capturing temporal dynamics with fewer parameters than other recurrent networks, and a Shallow Decoder Network (SDN) for low-cost state reconstruction. By leveraging this efficient architecture, the model can incorporate physical couplings between fields and key operational parameters to accurately reconstruct the full-field temperature and fluid dynamic states from sparse temperature measurements. This design enhances the model's adaptability under varying conditions, making it suitable for building digital twin models of high-voltage apparatuses like GIL.

II. Method

A. Problem Setup

The overall modeling process can be framed as solving a dynamic field reconstruction inverse problem [13]. Generally, the forward problem uses a physical model $\mathcal{F}^\dagger$ to compute the full physical field $\Phi(x,t)$ from a given set of system parameters μ:

$$\mathcal{F}^\dagger : \mu \mapsto \Phi(x,t) \tag{1}$$

Conversely, the inverse problem aims to establish a mapping $\mathcal{G}^\dagger$ that estimates the full field $\Phi(x,t)$ from time-series sensor measurements $\{y_k\}_{k=t-L}^{t}$ and operating parameters $\{\mu_k\}_{k=t-L}^{t}$:

$$\mathcal{G}^\dagger : (\{y_k\}_{k=t-L}^{t}, \{\mu_k\}_{k=t-L}^{t}) \mapsto \Phi(x,t) \tag{2}$$

*Corresponding author: Ziqi Zhou

979-8-3315-8138-1

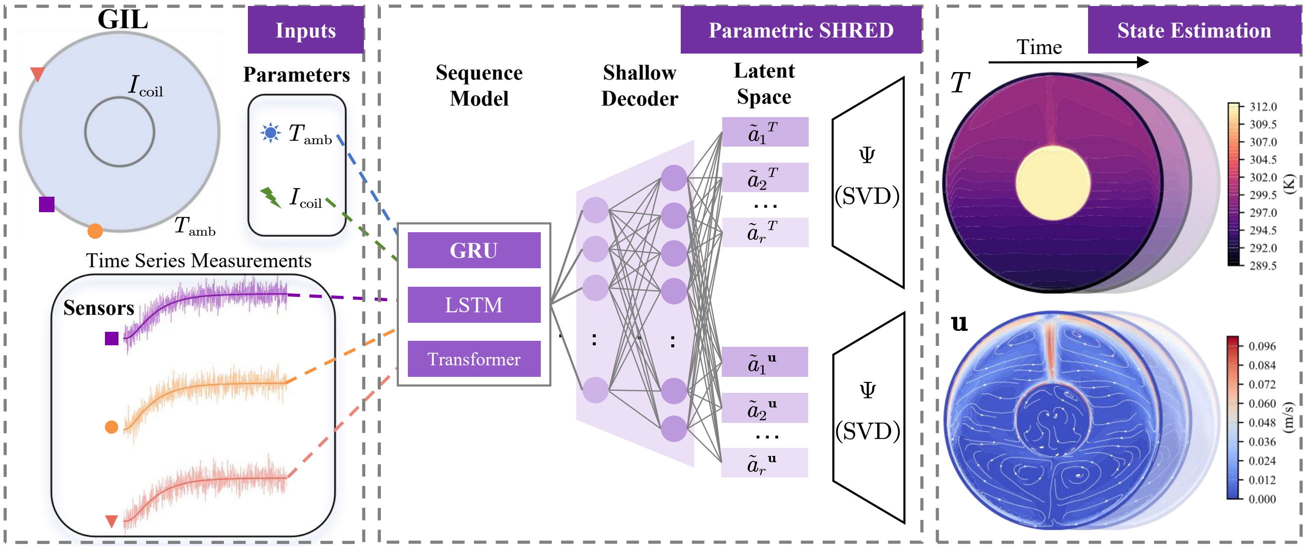

Fig. 1. Parametric SHRED architecture for real-time multi-physics field reconstruction in a GIL. The model takes sparse time-series sensor measurements and parameters as input and outputs the reconstructed temperature T and velocity $\mathbf{u}$ fields.

The objective of this work is to apply a parameterized surrogate model f_{SHRED} that efficiently and accurately approximates this complex inverse mapping $\mathcal{G}^{\dagger}$. The core of this framework is the parametric SHRED model, which integrates three key technologies: Proper Orthogonal Decomposition (POD), a parameter-aware input structure, and a recurrent decoder architecture. Fig. 1 illustrates the basic architecture of the model studied in this work. Based on full-order model simulation data of the GIL, a set of time-series sensor measurements and initial parameters can be obtained. After processing by the parameterized SHRED model, real-time state estimation of the temperature field T and velocity field $\mathbf{u}$ can be achieved.

B. Data Preparation and Formulation

The model input is built to encapsulate both temporal information and the physical operating state. For a given physical field of interest, denoted by Φ, a full-field simulation snapshot is represented as $\mathbf{X}^{\Phi,\mu_i} \in \mathbb{R}^{N_t \times N_s}$, where N_t is the number of time steps and N_s is the number of spatial grid points. To ensure the model generalizes across different conditions, data is generated for N_{traj} distinct trajectories, each defined by a set of physical parameters $\mu_i = \{\mu_{i,1}, \ldots, \mu_{i,N_p}\}$. At each time t, the input vector $x_{i,t}$ is constructed by combining the sparse sensor readings $y_{i,t} = \mathbf{C}_{\text{sens}}\mathbf{X}_t^{\Phi,\mu_i} + \epsilon_t \in \mathbb{R}^{N_{\text{sens}}}$, where ϵ_t represents measurement noise, $\mathbf{C}_{\text{sens}}$ is the measurement matrix, with the corresponding operating parameters μ_i. The combined vector $x_{i,t} = [y_{i,t}^{\top}, \mu_i^{\top}]^{\top} \in \mathbb{R}^{(N_{\text{sens}}+N_p)}$ is then structured into overlapping time sequences of length $L+1$, such as $\mathbf{X}_{\text{seq}} = (x_{t-L}, x_{t-L+1}, \ldots, x_t)$, which serve as the final input to the recurrent network.

For the model output, the prediction of a low-dimensional representation of the high-dimensional physical fields is considered to ensure efficiency. The full-field data for each physical quantity $\mathbf{X}^{\Phi}$ is first scaled and then compressed using POD. Through Singular Value Decomposition (SVD), $\mathbf{X}^{\Phi} = \mathbf{U}\Sigma\mathbf{V}^{\top} \approx \mathbf{U}_{r_\Phi}\Sigma_{r_\Phi}\mathbf{V}_{r_\Phi}^{\top}$, a low-rank approximation is obtained by retaining the first r_Φ most significant modes, which capture the system's essential dynamics. The neural network's prediction target is the set of temporal coefficients, $\mathbf{A}^{\Phi} = \mathbf{X}^{\Phi}\mathbf{U}_{r_\Phi} \in \mathbb{R}^{(N_{traj} \times N_t) \times r_\Phi}$. To facilitate multi-physics inversion, the standardized temporal coefficients for all relevant fields (temperature T, velocity u, v) are concatenated into a single target vector: $\mathbf{A}_{\text{concat}} = [\tilde{\mathbf{A}}^T, \tilde{\mathbf{A}}^u, \tilde{\mathbf{A}}^v] \in \mathbb{R}^{(N_{traj} \times N_t) \times (r_T + r_u + r_v)}$

C. Parametric SHRED Model Architecture and Training

The parametric SHRED model is structured as a composite function designed to map the input sequence directly to the predicted POD coefficients:

$$\hat{\mathbf{A}}_t = f_{\text{SHRED}}(\mathbf{X}_{\text{seq}}) = f_{\text{SDN}}(f_{\text{GRU}}(\mathbf{X}_{\text{seq}})) \tag{3}$$

The architecture consists of two main components: a GRU-based encoder and a shallow decoder.

The GRU network serves as the encoder, chosen for its ability to effectively capture temporal dependencies while maintaining computational efficiency. Compared to more complex architectures like Long Short Term Memory (LSTM) networks, the GRU uses a simpler gating mechanism to control the flow of information through time [14]. It iteratively updates its hidden state via the process $h_k = (1-z_k)\odot h_{k-1} + z_k \odot \tilde{h}_k$, where the update gate z_k and candidate state $\tilde{h}_k$ are computed from the current input x_k and the previous state h_{k-1}. This mechanism allows the GRU to effectively capture and compress the spatio-temporal dynamics of the system into the final hidden state vector $h_t = f_{\text{GRU}}(\mathbf{X}_{\text{seq}}) \in \mathbb{R}^{N_{\text{hidden}}}$.

h_t is then passed to the SDN, a feed-forward multi-layer perceptron (MLP). The SDN decodes the features extracted by the GRU to estimate the target POD coefficients. It performs a mapping using 2 fully-connected layers, and the output is the

979-8-3315-8138-1

final predicted vector of POD coefficients $\hat{\mathbf{A}}_t = f_{\text{SDN}}(h_t) \in \mathbb{R}^{r_T+r_u+r_v}$.

The model's parameters are optimized by minimizing the difference between the predicted and true coefficients on the training dataset. By repeating this process over the whole dataset, the SHRED model learns the precise mapping from the sparse, parameter-aware time-series input to the low-dimensional representation of the high-dimensional physical fields. These coefficients are used to reconstruct the full high-dimensional fields, which is achieved by multiplying the coefficients for each field by the transpose of its corresponding POD basis matrix:

$$\hat{\mathbf{X}}_t^{\Phi} = \hat{\mathbf{A}}_t^{\Phi}\mathbf{\Psi}_{\Phi}^{\top} \in \mathbb{R}^{1\times N_s}, t = 1, 2, \dots, N_t \tag{4}$$

III. Full-order Modeling of GIL

This work focuses on a specific 550 kV GIL with a rated current capacity of 8 kA. To reduce computation time, this work uses a two-dimensional model, shown in Fig. 2, which includes the key dimensions and thermal processes.

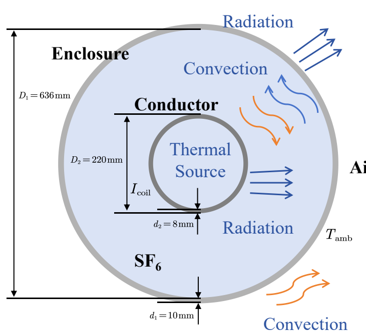

Fig. 2. Illustration of the schematic configuration for the Full-Order Modeling (FOM).

The temperature distribution inside the GIL is a result of a complex interplay between electromagnetic heating, fluid flow, and multiple modes of heat transfer. The main source of heat is Joule heating, caused by the load current I_{coil} flowing through the conductor. The heat transfer process in the GIL includes thermal conduction, convection, and radiation. In high-capacity GIL systems, the high temperature rise induced by high currents leads to strong natural convection within the insulating gas. To ensure efficient heat transfer, the gas flow is often in a turbulent state. Consequently, the fluid dynamics within the gas domain are described by the Reynolds-Averaged Navier-Stokes (RANS) equations with the Shear Stress Transport (SST) $k-\omega$ turbulence model. The SST model is well-suited for this application as it accurately captures flow characteristics both near the walls and in the free-stream region, making it robust for internal flow simulations. The heat exchange between the outer surface of the GIL enclosure and the surrounding ambient air is modeled as a convective boundary condition, which is implemented by using a heat transfer coefficient h_{air} and setting the external environmental temperature to the ambient temperature T_{amb}. The specific simulation settings can be found in reference [15].

By varying the two key input parameters—with the ambient temperature T_{amb} ranging from 10°C to 40°C and the load current I_{coil} from 1000A to 8800A, the full-order model was used to generate a total of 25 distinct simulation trajectories. Each trajectory representing an operational scenario was simulated for $N_t = 999$ time steps, with each step representing an interval of 30 seconds. The full-field solution at each time step comprises $N_s = 8,983$ spatial degrees of freedom, capturing the detailed temperature and velocity distributions. All simulations were carried out on a desktop equipped with an Intel Core Ultra 7 265KF CPU.

IV. Numerical Results and Discussion

To construct a low-dimensional representation of the system, a randomized SVD was first employed on the snapshot matrices for the temperature T, horizontal velocity u, and vertical velocity v fields. This process extracted a latent space spanned by the principal POD modes. As shown in Fig. 3, (a) illustrates the decay of singular values for each field, while (b) visualize the spatial patterns Ψ_{Φ} of the first three dominant POD modes. To preserve over 99.9% of the total energy in the latent space, the dimensionality was set to $r_T = 3$ for the temperature field and $r_u = 10$ and $r_v = 10$ for the velocity components.

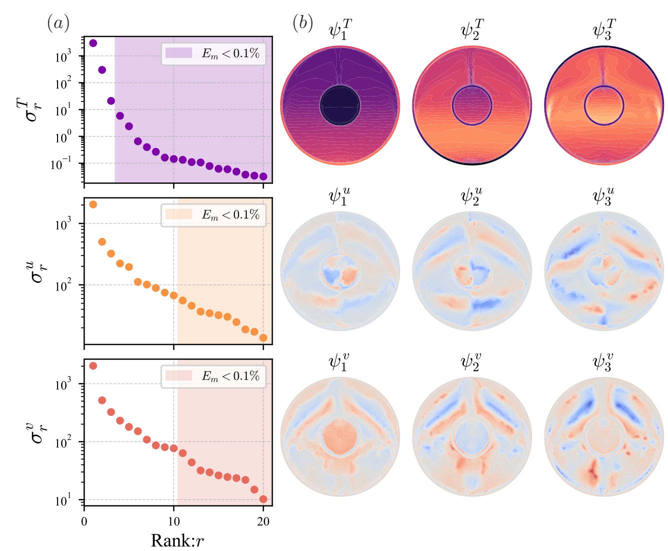

Fig. 3. Decay of singular values and the first 3 POD modes Ψ^{Φ} for the temperature field and velocity components.

With the dominant POD modes identified, the sensor locations for the model input were determined. For this study, 3 sensor locations were selected randomly on the GIL's outer enclosure. The measurements from these randomly placed sensors formed the sparse input for the model.

A. Model Performance and Reconstruction Accuracy

For the model training and evaluation, the 25 distinct simulation trajectories were divided into three sets: 18 trajectories

979-8-3315-8138-1

were used for training, 5 for validation, and 2 for testing. To assess the model's performance under realistic conditions, its robustness to sensor noise was evaluated. The noise $\epsilon = \mathcal{N}(0,1) \cdot \sigma_{data} \cdot l$ was added to the sensor measurements y at each time step, scaled relative to the data's natural variance. $\mathcal{N}(0,1)$ is standard Gaussian noise, σ_{data} is the standard deviation of the sensor training data, and l is the variable noise level. To ensure statistical reliability, 10 independent experiments were performed for each noise level.

Fig. 4 shows the predicted temporal evolution of the first three temperature POD coefficients $(\tilde{a}_1^T, \tilde{a}_2^T, \tilde{a}_3^T)$ under noisy conditions. The mean prediction (solid lines) accurately tracks the ground truth from the full-order model (dashed lines), even at high noise levels. The narrow 95% confidence interval (shaded region) further confirms the model's stability and its ability to effectively filter measurement noise while capturing the underlying system dynamics.

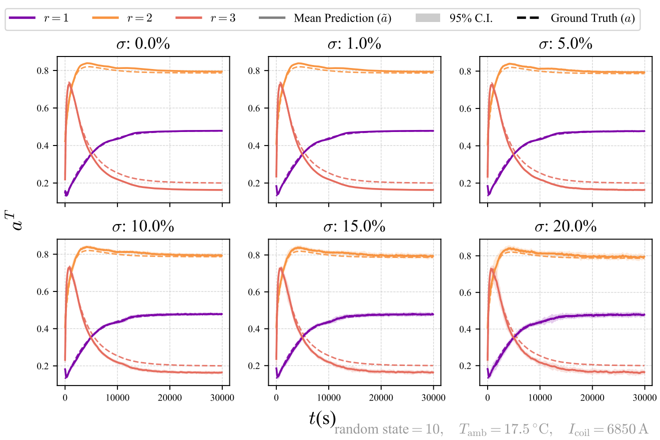

Fig. 4. Comparison of the parametric SHRED output to the Ground Truth for the first three temperature POD coefficients $(\tilde{a}_1^T, \tilde{a}_2^T, \tilde{a}_3^T)$.

The final step is the reconstruction of the high-dimensional physical fields, achieved by combining the predicted temporal coefficients with their corresponding spatial modes. Fig. 5 provides a visual comparison of the reconstructed fields against the ground truth for a representative time instance from the test set. The figure is organized into three panels: (a) displays the temperature T and velocity $\mathbf{u}$ fields as predicted by the SHRED model, (b) shows the corresponding ground truth fields from the full-order model simulation, and (c) illustrates the point-wise absolute error between the prediction and the truth, confirming the high accuracy of the reconstruction. Besides, the model exhibit good computational efficiency. The average prediction time for all time steps was measured to be around 0.2s, which satisfies the requirements for real-time digital twin applications.

B. Generalization and Stability Analysis

To ensure the model's reliability, two stability analyses were performed to test its sensitivity to both the data partitioning and the specific placement of sensors.

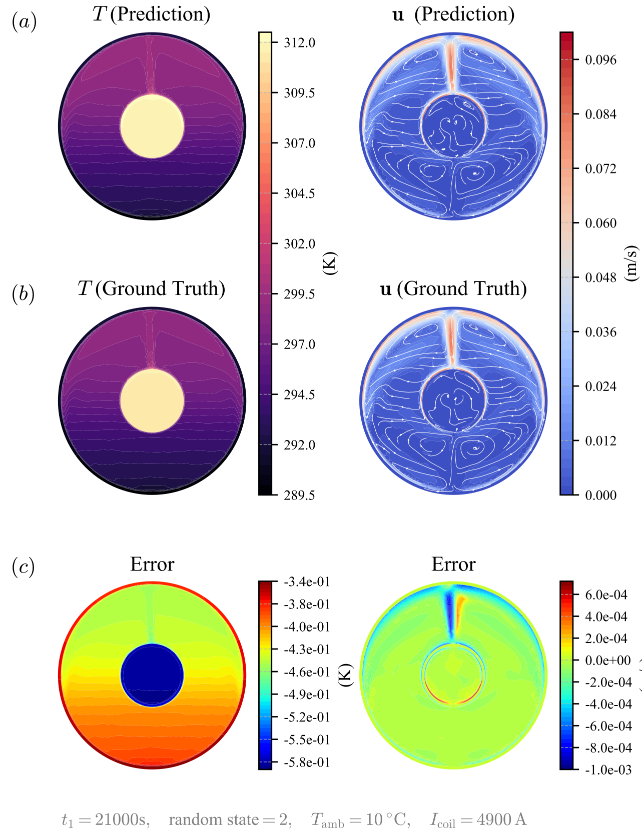

Fig. 5. Comparison of reconstructed fields T and $\mathbf{u}$: (a) SHRED model prediction, (b) ground truth from simulation, and (c) the point-wise absolute error.

First, the 25 trajectories were randomly split into training, validation, and testing sets 15 different times, and the model was retrained for each split. As shown in Fig. 6, the resulting temperature Mean Absolute Percentage Error (MAPE) shows low variation across the 15 trials, indicating that the model's performance is not dependent on a specific data partition.

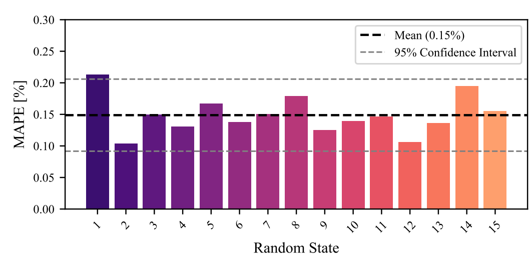

Fig. 6. Model stability analysis across 15 different data splits.

Second, the model was retrained using 10 different random sensor configurations, as shown in Fig. 7. To isolate the effect of sensor placement, the data partition (the division of the 25 trajectories into training, validation, and test sets) was kept identical for all 10 trials. The model was then retrained and evaluated for each of the 10 unique sensor layouts.

979-8-3315-8138-1

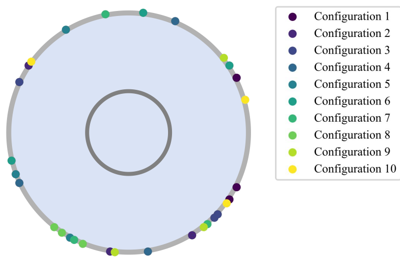

Fig. 7. The 10 random sensor configurations used for stability testing.

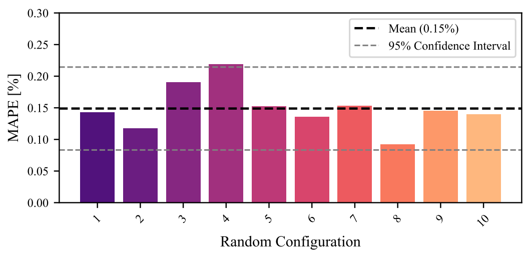

Fig. 8. Temperature MAPE results for the 10 sensor configurations.

The results are presented in Fig. 8, which shows the distribution of the temperature MAPE for the 10 different sensor configurations. The plot clearly shows a low variance in performance across the different layouts. The narrow 95% confidence interval provides evidence that the specific random placement of sensors has a negligible impact on the final prediction accuracy of the temperature field.

C. Ablation Study of Sequence Model Architecture

The GRU was initially chosen for its balance of efficiency and performance. To validate this choice, an ablation study was conducted by replacing the GRU component with the SHRED architecture with two other prominent sequence models: the LSTM network and a standard Transformer encoder. The results of this comparison are summarized in Table I. The GRU-based model not only achieves the lowest MAPE but also boasts the fastest training time.

TABLE I
PERFORMANCE COMPARISON OF DIFFERENT SEQUENCE MODELS

Sequence Model	**Training Time (s)**	**MAPE (%)**
GRU (Proposed)	20.1	0.11
LSTM	22.0	0.14
Transformer	40.1	0.51

V. CONCULSION

This work applies a parametric SHRED architecture for adaptable multi-physics field reconstruction in a 550kV GIL. Based on only 3 sparse temperature sensors and key operational parameters, the model was trained to estimate the multi-physics field. By learning the spatiotemporal relationships embedded in the sparse data, the parametric SHRED model accurately reconstructs high-dimensional fields with a MAPE below 0.25%, even under noisy conditions.

The model's stability against sensor noise and placement, and its ability to adapt to different scenarios, makes it suitable for real-world DT applications. This makes the parametric SHRED a powerful tool for building efficient and adaptable DTs for non-intrusive, real-time state estimation in high-voltage apparatus. Future work will extend this parametric approach to more complex equipment and a broader range of operating conditions.

REFERENCES

[1] Milan Groshev, Carlos Guimaraes, Jorge Martín-Pérez, and Antonio de la Oliva. Toward intelligent cyber-physical systems: Digital twin meets artificial intelligence. *IEEE Communications Magazine*, 59(8):14–20, 2021.

[2] Alexandra I Khalyasmaa, Alina I Stepanova, Stanislav A Eroshenko, and Pavel V Matrenin. Review of the digital twin technology applications for electrical equipment lifecycle management. *Mathematics*, 11(6):1315, 2023.

[3] Hermann J Koch. *Gas insulated transmission lines (GIL)*. John Wiley & Sons, 2011.

[4] Joong Kyoung Kim, Sung Chin Hahn, Kyong Yop Park, Hong Kyu Kim, and Yeon Ho Oh. Temperature rise prediction of ehv gis bus bar by coupled magnetothermal finite element method. *IEEE Transactions on Magnetics*, 41(5):1636–1639, 2005.

[5] Wei Du, Zikun Zhao, Xiaolong Li, Chen Cao, and Wen Wang. Simulation of three-dimensional temperature distribution for gil insulator based on fvm. In *2020 IEEE 1st China International Youth Conference on Electrical Engineering (CIYCEE)*, pages 1–5. IEEE, 2020.

[6] Shucan Cheng, Yanpu Zhao, Kejia Xie, Bin Hu, Jinxian Zhang, and Xingxiong Yang. A novel method for fast computation of the temperature rise and optimal design of gil based on thermal network model. *Energy*, 289:130034, 2024.

[7] Zheming Tong and Yue Li. Real-time reconstruction of contaminant dispersion from sparse sensor observations with gappy pod method. *Energies*, 13(8):1956, 2020.

[8] Marco Baldan, Paolo Di Barba, and David A Lowther. Physics-informed neural networks for inverse electromagnetic problems. *IEEE Transactions on Magnetics*, 59(5):1–5, 2023.

[9] Jan P Williams, Olivia Zahn, and J Nathan Kutz. Sensing with shallow recurrent decoder networks. *Proceedings of the Royal Society A*, 480(2298):20240054, 2024.

[10] Stefano Riva, Carolina Introini, Antonio Cammi, and J. Nathan Kutz. Robust state estimation from partial out-core measurements with shallow recurrent decoder for nuclear reactors. *Progress in Nuclear Energy*, 189:105928, 2025.

[11] Stefano Riva, Carolina Introini, J Nathan Kutz, and Antonio Cammi. Towards efficient parametric state estimation in circulating fuel reactors with shallow recurrent decoder networks. *arXiv preprint arXiv:2503.08904*, 2025.

[12] Ziqi Zhou, Bo Liu, Wenchuan Chen, Shaoshuai Su, Ziwei Zhang, Qiang Sun, and Wensheng Gao. Robust multi-physics field reconstruction via sparse enclosure temperature measurements for gil digital twin application. *International Journal of Electrical Power & Energy Systems*, 172:111273, 2025.

[13] Steven L Brunton, Bernd R Noack, and Petros Koumoutsakos. Machine learning for fluid mechanics. *Annual review of fluid mechanics*, 52(1):477–508, 2020.

[14] Farhad Mortezapour Shiri, Thinagaran Perumal, Norwati Mustapha, and Raihani Mohamed. A comprehensive overview and comparative analysis on deep learning models: Cnn, rnn, lstm, gru. *arXiv preprint arXiv:2305.17473*, 2023.

[15] Mahesh T Dhotre, Jakub Korbel, Xiangyang Ye, Joerg Ostrowski, Sami Kotilainen, and Martin Kriegel. Cfd simulation of temperature rise in high-voltage circuit breakers. *IEEE Transactions on Power Delivery*, 32(6):2530–2536, 2017.

979-8-3315-8138-1