FROM SYNTHESIS TO KINETICS: A DATA-DRIVEN DEEP LEARNING FRAMEWORK FOR PROCESS-AWARE FERROELECTRIC DYNAMICS

Anonymous authors
Paper under double-blind review

ABSTRACT

Physical modeling of emerging ferroelectric $Hf_xZr_{1-x}O_2$ (HZO) materials is constrained by incomplete theoretical descriptions of domain kinetics and the difficulty of formulating efficient equations for complex evolutionary mechanisms, such as the wake-up effect. To address this challenge, we propose a data-driven Deep Learning framework utilizing an optimized Residual Artificial Neural Network (ResANN). By incorporating critical fabrication parameters as direct inputs, our model learns the intrinsic process-property mappings directly from experimental data. While standard ML-based approaches often leverage idealized simulation data, this work expands the scope by rigorously training on large-scale device measurement data, thus establishing a direct bridge between fabrication parameters and material behavior. By leveraging an optimized deep learning framework and Transfer Learning, our methodology demonstrates exceptional capability in capturing highly non-linear polarization dynamics. We show that this approach yields superior accuracy (Adjusted $R^2 \approx 0.998$) compared to traditional physical equations, effectively predicting hysteresis evolution and achieving zero-shot generalization to untrained material thicknesses. Consequently, this framework serves as a predictive tool for inverse material design, guiding the optimization of synthesis protocols for next-generation ferroelectrics materials.

1 INTRODUCTION

Establishing accurate physical descriptions for nanoscale ferroelectric films is increasingly challenged by the stochastic nature of intrinsic material kinetics [1]. Traditional look-up table (LUT) methods often struggle to capture continuous, multi-dimensional property evolution, leading to convergence issues in system-level analysis [2, 3]. While Deep Learning (DL) offers high-dimensional feature extraction capabilities to identify complex process-property mappings [4, 5], state-of-the-art approaches primarily rely on idealized TCAD simulations. This creates a "reality gap" as simulations fail to capture real-world process variations and defect dynamics [6]. Therefore, a robust characterization framework must be driven by experimental measurement data.

Ferroelectric (FE) materials have attracted significant attention for next-generation non-volatile memory (NVM) technologies due to their distinct polarization switching dynamics and high compatibility with CMOS processes [7, 8]. At the core of these technologies lie FE capacitors (FeCAPs), typically based on polycrystalline films (e.g., HfZrOx). However, accurately describing the physics of these FE layers faces substantial challenges, particularly in capturing dynamic evolutionary behaviors such as the *wake-up* effect—a phase transition and domain depinning process driven by electric field cycling and defect redistribution [6]. To address these challenges, it is essential to develop high-precision modeling approaches capable of capturing intricate material kinetics across varying fabrication parameters.

In this work, we present a novel measurement-driven Residual Artificial Neural Network (ResANN) framework designed to learn the intrinsic process-property mappings governing FeCAP polarization kinetics. By deploying an optimized residual topology combined with Transfer Learning, our approach not only captures complex non-linear evolutionary dynamics but also robustly extrapolates to uncharacterized material geometries with minimal data dependency. Our main contributions are:

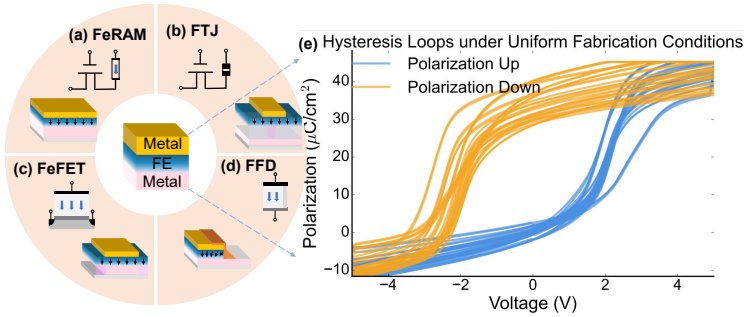


Figure 1: The hierarchy of ferroelectric memory devices anchored by the fundamental Metal-Ferroelectric-Metal (MFM) stack: (a) FeRAM, (b) FTJ, (c) FeFET, and (d) FFD. (e) Experimental characterization of polarization kinetics in polycrystalline HZO films, illustrating the non-linear hysteresis evolution and domain depinning during the wake-up process.

- We propose a novel methodology to develop a measurement-driven ResANN model for FeCAPs. This model exhibits remarkable accuracy in learning the intrinsic *process-property* mapping from real fabrication data, effectively bridging the gap between material synthesis and physical behavioral prediction.
- The framework demonstrates powerful zero-shot generalization capabilities by accurately extrapolating material performance to unseen geometric configurations (e.g., untrained film thicknesses), proving that the model has learned the underlying physical scaling laws.
- By leveraging Transfer Learning, the fine-tuned model can rapidly adapt to batch-specific process variations with a minimal amount of data, demonstrating the efficiency and robustness of our AI-driven approach in accelerating material optimization.

To accelerate ferroelectric material discovery and process optimization, the complete dataset and source code will be open-sourced upon acceptance.

2 Background

2.1 Ferroelectric Material Physics

The Ferroelectric Capacitor (FeCAP) serves as the fundamental building block for characterizing the intrinsic polarization behavior of ferroelectric materials, replacing the conventional linear dielectric with a ferroelectric film (e.g., doped HfO_2) [9]. Understanding the physics of the FeCAP is a prerequisite for optimizing all derived FE material technologies. Fig. 1 illustrates the FeCAP structure, consisting of a ferroelectric layer sandwiched between two metal electrodes. This metal-ferroelectric-metal (MFM) stack forms the basis of FeRAM (Fig. 1a), where the FeCAP is integrated with a selection transistor for charge storage [10]. Similarly, in the Ferroelectric Tunnel Junction (FTJ) (Fig. 1b), the FeCAP structure utilizes an ultra-thin FE layer to modulate tunneling resistance [11]. In the Ferroelectric Field-Effect Transistor (FeFET) (Fig. 1c), the FeCAP is integrated directly into the gate stack, coupling the ferroelectric polarization with the semiconductor channel conductance [12]. Recently, novel architectures like the Ferroelectric Fin Diode (FFD) (Fig. 1d) have further diversified the application of this material system [13]. Therefore, establishing a high-fidelity model for the FeCAP is critical, as it bridges the gap between material synthesis and the functional design of complex memory devices.

The core physical characteristic of the FE layer is its non-linear hysteretic relationship between polarization and electric field (P-V), as shown in Fig. 1e. Binary information is stored as distinct remanent polarization states ($+P_R$ and $-P_R$) [14]. However, polycrystalline ferroelectric films, such as HZO, exhibit complex time-dependent evolutionary behaviors. Most notably, pristine devices often require a *wake-up* process—a cycling-induced phase transformation and redistribution of oxygen vacancies—to achieve stable ferroelectricity [15, 16]. As illustrated in Fig. 1e, the hysteresis loop opens up and the distinctness of the polarization states improves significantly after cycling.

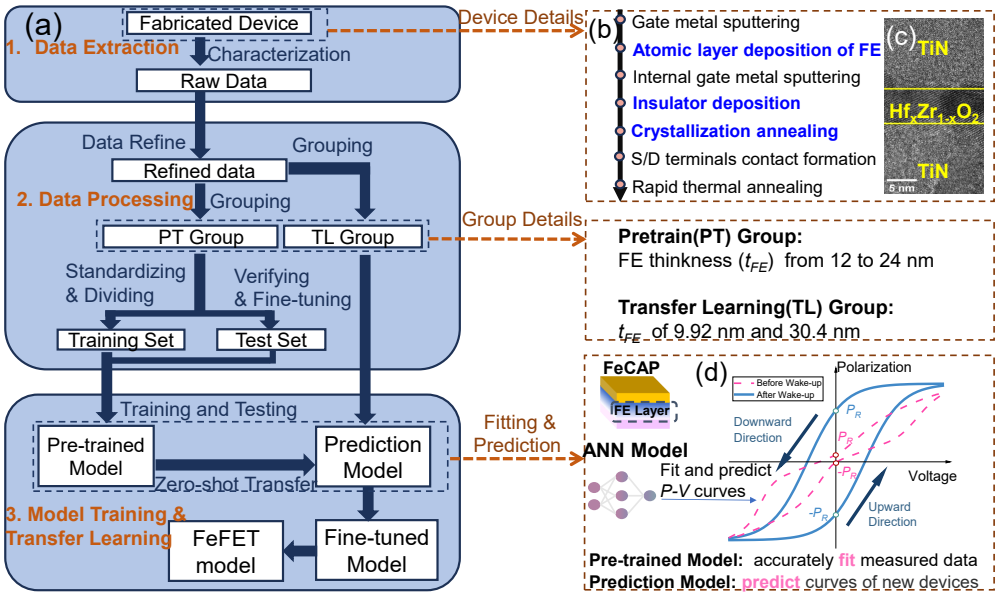


Figure 2: Overview of the proposed data-driven modeling framework: from material fabrication and characterization to deep learning-based kinetic prediction.

Accurately modeling this dynamic evolution of P_R is notoriously difficult for traditional physical equations due to the stochastic nature of domain nucleation and growth [9].

2.2 Deep Learning for Material Modeling

Machine Learning (ML), and specifically Deep Learning (DL), represents a paradigm shift in modeling complex physical systems where theoretical descriptions are incomplete or computationally prohibitive [17]. Unlike traditional regression, Artificial Neural Networks (ANNs) function as universal approximators capable of extracting high-dimensional features from raw material data. Through supervised learning, ANNs can deconvolute the intricate, non-linear mappings between fabrication parameters (inputs) and functional material properties (outputs) [18]. This capability makes them uniquely consistent in handling the subtle variations in device performance caused by process fluctuations [19–21].

Furthermore, Transfer Learning (TL) provides a sophisticated mechanism to generalize learned physical laws across different material domains [22, 23]. In the context of material modeling, obtaining exhaustive datasets for every possible film thickness or annealing condition is experimentally expensive. TL allows a model pre-trained on a source domain (e.g., a specific thickness range) to transfer its latent knowledge of polarization physics to a target domain with sparse data. By fine-tuning the network weights, the model adapts rapidly to new material configurations, significantly accelerating the exploration of the design space [6].

2.3 Related Work and Motivation

Existing modeling methodologies typically rely on synthetic data generated from idealized physical equations or TCAD simulations [24]. However, emerging ferroelectric materials exhibit complex kinetic behaviors—such as grain boundary effects and interface trap dynamics—that are not fully captured by current theoretical models [9]. Consequently, models trained on synthetic data often suffer from a ”reality gap,” failing to accurately reflect the electrical performance of fabricated devices. While some ML studies have investigated ferroelectrics, they have mostly focused on atomic-scale density functional theory (DFT) calculations [8] or limited parameter extraction [6], lacking a direct link to macroscopic device behavior like the wake-up effect.

Moreover, traditional ML models for semiconductor devices are often ”single-instance” solutions, designed for a specific device configuration [24–28]. Any modification to material parameters (e.g., changing the HZO film thickness) typically necessitates a complete retraining process, which is com-

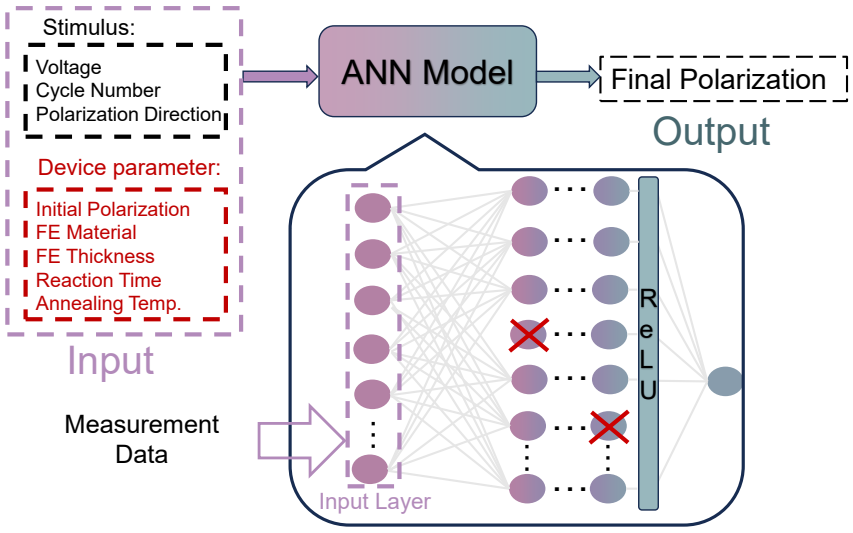


Figure 3: Schematic of the process-aware ResANN architecture. By explicitly integrating multidimensional synthesis parameters with dynamic stimuli, the network learns to map the intrinsic process-property relationships to non-linear polarization kinetics.

putationally inefficient. To address these limitations, our work introduces a robust, measurement-driven deep learning framework. By training directly on large-scale experimental characterization data, our model captures the true phenomenological behavior of the material. Furthermore, we integrate transfer learning to endow the model with strong generalization capabilities, enabling it to predict the polarization dynamics of new material structures without learning from scratch.

3 Machine Learning-Based Model Methodology

Fig. 2 illustrates our proposed end-to-end framework for kinetic modeling. This methodology integrates three synergistic components: precision material fabrication and electrical characterization, rigorous data refinement for physical consistency, and deep learning-based kinetic modeling.

3.1 Material Synthesis and Electrical Characterization

In this work, we synthesized and characterized ferroelectric capacitors (FeCAPs) utilizing $Hf_xZr_{1-x}O_2$ (HZO, where x=0, 0.2, or 0.5) as the functional layer. The material stack was fabricated on a highly doped *Si* wafer. After surface cleaning, a $30\,\mathrm{nm}$ *TiN* bottom electrode was deposited via reactive magnetron sputtering. Subsequently, the ferroelectric $Hf_xZr_{1-x}O_2$ film was grown using atomic layer deposition (ALD), allowing for precise atomic-scale control over stoichiometry and thickness. Top electrodes consisting of $30\,\mathrm{nm}$ *TiN* and $30\,\mathrm{nm}$ *W* were deposited by sputtering and patterned through standard photolithography and lift-off processes. Finally, to induce crystallization and stabilize the orthorhombic ferroelectric phase, the capacitors underwent post-metallization annealing in an N_2 ambient.

Electrical characterization was performed using a Keithley 4200A-SCS parameter analyzer. To capture the intrinsic polarization switching dynamics, P–V hysteresis loops were measured using a triangular voltage excitation ($0 \rightarrow +\textit{Cycle voltage} \rightarrow -\textit{Cycle voltage} \rightarrow 0$). The *polarization direction* switches in response to the electric field vector. To systematically characterize the time-dependent *wake-up* dynamics and domain evolution, tests were conducted at logarithmically spaced *cycle numbers* (e.g., $10^2, 10^3, 10^4$). This comprehensive characterization campaign resulted in a total of *313 measurements* across *128 distinct devices*, providing a rich dataset linking process variations to kinetic behavior.

3.2 Data Refinement and Feature Engineering

Following electrical characterization, we implemented a rigorous data processing pipeline to ensure the physical integrity and reliability of the dataset. This pipeline consists of several key steps: (1) Breakdown Filtration: Excluding devices exhibiting hard breakdown or excessive leakage currents

that obscure ferroelectric switching. (2) Physical Consistency Check: Assessing the quality of hysteresis loops to ensure they reflect genuine ferroelectric switching mechanisms rather than parasitic effects. (3) Outlier Suppression: Filtering extreme magnitude anomalies. After these refinement steps, the high-quality dataset comprised 100 devices and approximately 650,000 sampling points. This meticulous curation ensures that the subsequent Deep Learning model is trained on robust, representative data capturing the true material physics.

To facilitate robust model training, devices were categorized into distinct groups based on fabrication parameters (as shown in Fig. 2). We ensured balanced sample sizes across groups to prevent bias. Crucially, to test the model's ability to generalize physical laws to unseen geometries, we excluded devices with FE thicknesses of $9.92\,\mathrm{nm}$ and $30.4\,\mathrm{nm}$ from the training set. These specific groups were reserved strictly for Transfer Learning (TL) validation. Standardization was applied to all input features to normalize the dynamic ranges of different physical quantities. Finally, the dataset was partitioned into training and testing sets at the device level (see Fig. 2). This device-level splitting is critical to prevent data leakage, ensuring the model is evaluated on its ability to predict the behavior of entirely new physical devices rather than memorized points.

3.3 Deep Learning Framework

The core of our methodology involves training a high-capacity Artificial Neural Network (ANN) to learn the latent process-property mappings. To determine the optimal network topology that balances model complexity with generalization power, we employed Neural Architecture Search (NAS) using Neural Network Intelligence [29]. This automated search process explored a vast space of hyperparameters and architectural configurations on the validation subset.

The optimization objective was to minimize the Mean Squared Error (MSE). To comprehensively evaluate the capture of material physics, we also monitored Adjusted R^2, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) [30]. The Adjusted R^2 is particularly critical as it quantifies the variance in polarization kinetics explained by the model features.

Upon training a high-precision base model, we leveraged Transfer Learning (TL) to extend the model's applicability to the reserved device groups (unseen thicknesses). By freezing the early layers that capture fundamental switching physics and fine-tuning only specific weights, the model adapts to new material dimensions with minimal data. This capability demonstrates that the ANN has learned universal physical principles rather than simply memorizing the training set. Robustness was further verified via cross-validation, systematically rotating the hold-out groups to ensure stability across all material configurations. Finally, the optimized ANN was encapsulated into a Verilog-A model for SPICE co-simulation.

3.4 Optimized Neural Architecture

The deployed model, illustrated in Fig. 3, is a Residual Artificial Neural Network (ANN) specifically engineered to approximate the high-dimensional, non-linear mapping connecting material synthesis parameters to polarization kinetics.

The input vector combines dynamic measurement features (Voltage, Cycle number) and static process parameters (FE thickness, Reaction time, Annealing Temperature). This hybrid input allows the model to learn how specific fabrication conditions influence the dynamic wake-up process.

As shown in Fig. 3, the architecture comprises: (1) a Feature Embedding Layer (256 neurons) to capture feature interactions; (2) three Residual Blocks, each containing two dense layers and a skip connection ($y = F(x) + x$) to mitigate gradient vanishing and learn incremental polarization changes; and (3) a Regression Head to map features to the scalar output. We utilized Leaky ReLU activation, Dropout (rate=0.05), and the Adam optimizer with early stopping.

We utilized the Leaky Rectified Linear Unit (Leaky ReLU) activation function [31] to ensure robust gradient flow. To prevent overfitting, we applied Dropout regularization (rate=0.05) within each residual block. The network was optimized using the Adam algorithm [32, 33], incorporating an early stopping mechanism that halts training if the validation loss plateaus for 50 epochs [34]. This residual design ensures the model is both accurate and stable.

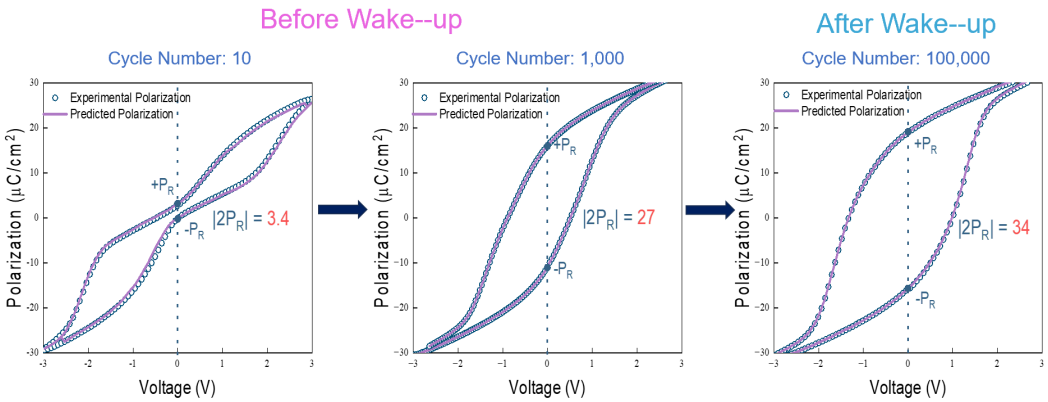


Figure 4: Model evaluation against temporal material dynamics: The ResANN accurately reconstructs the progressive expansion of P-V hysteresis loops, capturing the non-linear saturation of remanent polarization (P_R) throughout the wake-up cycling regime.

Table 1: Comparison with different models.

	MSE	RMSE	MAE	Adj. R^2
Physical Equation-based Model [35]	**14.866**	**3.856**	**1.971**	**/**
Random Forest Model	31.026	5.570	4.094	0.898
Support Vector Model	17.110	4.136	2.681	0.944
LASSO Model	81.024	9.001	6.142	0.814
Machine Learning-based Model	**0.724**	**0.851**	**0.343**	**0.998**

4 Result and Discussion

In this section, we evaluate the efficacy of the proposed data-driven framework in characterizing the complex polarization kinetics of HZO films. First, we present the performance of the pre-trained ResANN on the large-scale dataset to demonstrate its capability in capturing intrinsic material behaviors. The generated hysteresis loops are benchmarked against experimental measurements from the testing set. Subsequently, we assess the model's generalization power via transfer learning, specifically predicting the kinetic response of unseen ferroelectric layer thicknesses without prior exposure. Finally, the fine-tuned model is validated through SPICE co-simulation to verify its applicability in circuit-level design.

4.1 Evaluation of Kinetic Modeling Fidelity

As detailed in Subsection 3.2, the experimental dataset was systematically partitioned into 14 distinct process groups. To rigorously test the model's ability to extrapolate physical laws to new material geometries, the groups containing devices with FE thicknesses of $9.92\,\mathrm{nm}$ and $30.4\,\mathrm{nm}$ were excluded from the primary training phase. This yielded a robust training corpus of 76 devices (approx. 600,000 sampling points) across 10 groups, while 19 devices across 3 groups served as the unseen testing set.

The quantitative evaluation, summarized in Table 1, demonstrates the superior fidelity of our approach. After convergence, the ResANN achieved a Mean Squared Error (MSE) of **0.724** on the testing set. This exceptionally low error magnitude indicates the model's precision in tracking the highly non-linear polarization trajectories. Crucially, the model attained an *adjusted* R^2 of 0.998. This near-unity value signifies that $99.8\,\%$ of the variance in the polarization response—driven by complex domain nucleation and growth mechanisms—is effectively explained by the network's latent features. This confirms that the ResANN has successfully internalized the underlying physics governing the ferroelectric hysteresis relationships. To our knowledge, this represents the most accurate data-driven approach for modeling ferroelectric kinetics reported to date [6, 8, 36, 37].

To further contextualize the efficacy of our AI method, we benchmarked the ResANN against established baselines, including statistical regression models (LASSO) and traditional physics-based compact models (see Table 1). While regression models are often employed for simple parameter

Table 2: 14-fold Cross-Validation results.

Fold	MSE	MAE	Adj. R^2	Fold	MSE	MAE	Adj. R^2
1	0.499	0.187	0.999	2	1.188	0.604	0.986
3	0.508	0.245	0.999	4	0.791	0.322	0.998
5	2.195	0.571	0.971	6	0.899	0.369	0.989
7	0.580	0.349	0.999	8	0.243	0.795	0.996
9	0.450	0.209	0.999	10	0.494	0.229	0.999
11	0.422	0.151	0.999	12	0.319	0.193	0.999
13	0.335	0.242	0.997	14	0.484	0.151	0.999
Average	**0.672**	**0.330**	**0.995**				

extraction [38, 39], they struggle to capture the full hysteresis loop. For the physics-based benchmark, we selected the Nucleation-Limited Switching (NLS) model, widely regarded for its accuracy in describing polycrystalline ferroelectrics [35]. However, a critical limitation of the NLS model (and similar physical equations) is its static nature; it typically describes the stabilized post-wake-up phase and fails to account for the dynamic evolution of material properties during cycling. In contrast, our ResANN significantly outperforms all baselines, reducing the error by orders of magnitude. This advantage stems from the neural network's ability to approximate the high-dimensional transfer functions that link fabrication process variations to dynamic material performance.

To visualize this predictive capability, we randomly selected a representative set of P-V hysteresis loops from the testing set. As illustrated in Figure 4, the blue scatters denote the experimental measurements, while the solid lines represent the ResANN predictions. The model exhibits exceptional alignment with the experimental data, capturing subtle non-linearities. Most notably, the model accurately reproduces the wake-up effect—a phenomenon driven by the redistribution of oxygen vacancies and phase transformation during electric field cycling [16]. As observed, the remanent polarization (P_R) exhibits substantial growth as the cycle number increases from 10 to 100,000. Traditional models often treat parameters like P_R as constants or require complex empirical corrections to model this evolution [40–42]. Our results indicate that the ResANN has effectively internalized the intrinsic mechanisms of defect migration and domain depinning, enabling it to predict the full evolutionary trajectory of the ferroelectric film from its pristine state to saturation [16]. This capability is pivotal for diagnosing the kinetic stability of HZO stacks and identifying optimal fabrication windows to minimize wake-up latency [43]. Consequently, this precise modeling of process-dependent kinetics provides a powerful computational tool for accelerating the discovery of high-performance ferroelectric materials and guiding the optimization of next-generation synthesis protocols.

4.2 Generalization to Unseen Geometries via Transfer Learning

To rigorously evaluate the model's ability to extrapolate physical laws to uncharacterized material dimensions, we employed a Transfer Learning (TL) strategy. Specifically, we tasked the pre-trained ResANN to predict the kinetic response of FeCAPs with ferroelectric film thicknesses of $9.92\,\mathrm{nm}$ and $30.4\,\mathrm{nm}$—geometries that were deliberately excluded from the source training domain. This zero-shot prediction scenario tests whether the network has internalized the intrinsic scaling laws governing polarization density rather than merely memorizing specific fabrication batches. Following this initial assessment, we implemented a Domain Adaptation phase, where the model was fine-tuned on the target data. To guarantee statistical robustness and mitigate selection bias, a rigorous 14-fold cross-validation protocol was executed across the partitioned dataset.

Fig. 5 illustrates the zero-shot predictive performance of the base model. Even without prior exposure to these specific thicknesses, the ResANN exhibits a remarkable capacity to reconstruct the full hysteresis envelope. The comparison reveals that the model accurately captures the non-linear switching dynamics both before and after the wake-up phase. For the $9.92\,\mathrm{nm}$ capacitors, the model achieves high fidelity (high Adjusted R^2), attributed to the network's effective interpolation within the learned feature space. Crucially, for the $30.4\,\mathrm{nm}$ samples—which represent a significant deviation from the training distribution—the model maintains robust explanatory power with Adjusted R^2 values of 0.95 and 0.97. Minor deviations observed in the switching directionality are attributed to stochastic process variations inherent to the fabrication of thicker films, rather than model deficiency. These results substantiate that the ResANN has successfully disentangled the geometric

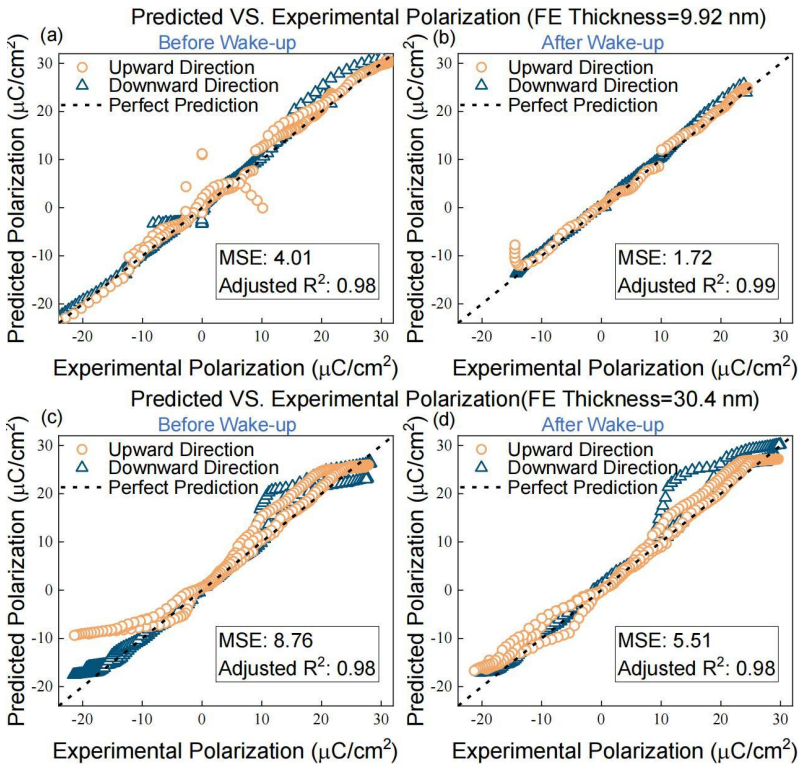


Figure 5: Evaluation of Zero-Shot Generalization: Comparison of experimental measurements vs. ResANN predictions for untrained FE film thicknesses ($9.92\,\mathrm{nm}$ and $30.4\,\mathrm{nm}$), demonstrating robust extrapolation of polarization kinetics.

scaling factors from the fundamental polarization mechanisms, satisfying a key criterion for interpretable scientific machine learning [44].

Subsequently, to optimize the model for specific device batches, we applied the fine-tuning procedure. The 14-fold cross-validation results, summarized in Table 2, validate the stability of this domain adaptation. By calibrating the high-level weights to the specific process group, the model achieves a minimized average MSE of **0.672**. This significant accuracy enhancement confirms that while the pre-trained model captures universal physical laws, the fine-tuning step effectively compensates for batch-to-batch process variability. The exceptional average Adjusted R^2 of 0.995 indicates that our framework provides a nearly complete statistical description of the polarization kinetics across the entire design space [45], proving its viability for rapid material screening and process optimization.

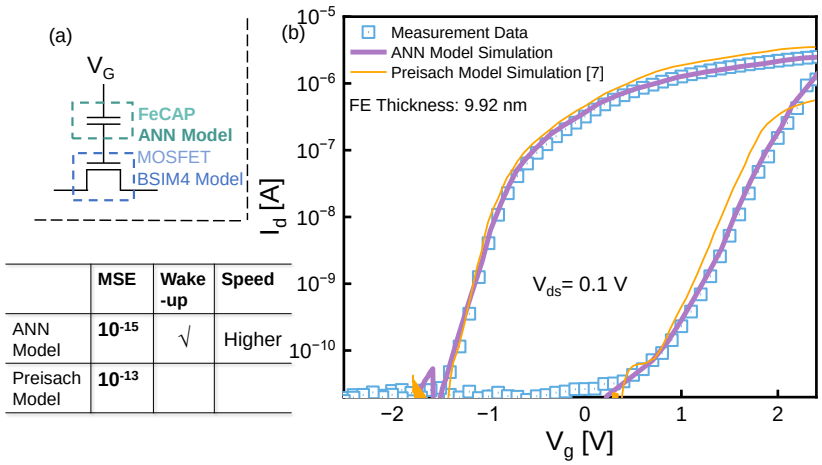
4.3 Validation of Material Kinetics via Device Integration

To verify our material model in a realistic device setting, we integrated the optimized ResANN into a Verilog-A block for system-level validation. We utilized the Ferroelectric Field-Effect Transistor (FeFET) as a physical test platform to check the consistency of our HZO model. Fundamentally, the FeFET places the ferroelectric capacitor (FeCAP) directly into the gate of a MOSFET (Fig. 1c), creating a system governed by the interaction between the material polarization and the semiconductor surface potential [40]:

$$P_{MOS} = P_{FE} \tag{1}$$

$$V_G = V_{FE} + V_{MOS} \tag{2}$$

Here, P_{FE} represents the polarization response of the HZO layer predicted by our ResANN, while V_{FE} denotes the voltage across the ferroelectric film. Eq. 1 and Eq. 2 ensure charge and voltage balance, linking the material's hysteretic state (Fig. 4) to the semiconductor charge transport (V_{MOS}, P_{MOS}), which we modeled using the standard BSIM4 model. Experimental validation was



	MSE	Wake-up	Speed
ANN Model	10^{-15}	√	Higher
Preisach Model	10^{-13}		

Figure 6: Validation of material kinetics via FeFET integration. The accurate prediction of transfer characteristics ($I_d - V_g$) for an uncharacterized process combination confirms the ResANN's ability to extrapolate intrinsic physical laws beyond the training domain.

conducted on FeFET devices fabricated with specific process conditions described in [46]. Fig. 6 presents the channel current obtained by sweeping the gate voltage. Crucially, the input configuration used in this validation represents a process combination not seen during training. Thus, the accurate prediction of the transfer characteristics (blue scatters) demonstrates that the ResANN has successfully extended the intrinsic material physics to a new device environment, rather than merely memorizing training patterns.

To further highlight the necessity of this data-driven approach, we compared our model with the traditional Preisach model, a widely used physics-based method [40, 41]. The simulation results (black line) reveal a critical limitation of traditional methods: while the Preisach model can approximate the static hysteresis loop, it is unable to capture the dynamic wake-up process caused by defect redistribution. In contrast, our ResANN (pink and purple lines) shows superior accuracy, reducing the MSE by two orders of magnitude compared to the physics-based baseline. This significant gap highlights that the neural network has learned the time-dependent kinetic laws of the material that static equations fail to describe. The successful reproduction of the binary states ('0' and '1') under a read voltage of $0.5\,\mathrm{V}$ confirms that the proposed framework provides a robust description of the ferroelectric layer. This establishes the ResANN as a critical tool for optimizing material synthesis, ensuring that complex kinetic behaviors are accurately accounted for in next-generation memory design.

5 Conclusion

This study establishes a novel data-driven framework for modeling the kinetic properties of ferroelectric HZO films, based on extensive real-device measurements. The optimized ResANN model achieves exceptional accuracy with an *adjusted* R^2 of 0.998, demonstrating its ability to precisely capture the complex, non-linear polarization evolution. Furthermore, by leveraging Neural Architecture Search and Transfer Learning, the framework exhibits strong generalization power. It effectively predicts the physical behavior of new material geometries (such as different film thicknesses) without needing to learn from scratch. Finally, the model was integrated into a device simulation platform, proving its practical ability to reproduce dynamic physical effects like the wake-up process in real-world test structures. This methodology bridges the gap between material synthesis and functional performance, providing a powerful tool to accelerate the material optimization and development of next-generation ferroelectric technologies.

REFERENCES

[1] F. Klemme et al., "Modeling emerging technologies using machine learning: Challenges and opportunities," in *2020 IEEE/ACM International Conference On Computer Aided Design (ICCAD)*, 2020, pp. 1–9.

[2] R. A. Thakker et al., "A novel table-based approach for design of finfet circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 28, no. 7, pp. 1061–1070, 2009.

[3] B. Chawla et al., "Motis-an mos timing simulator," *IEEE Transactions on Circuits and Systems*, vol. 22, no. 12, pp. 901–910, 1975.

[4] G. E. Karniadakis et al., "Physics-informed machine learning," *Nature Reviews Physics*, vol. 3, no. 6, pp. 422–440, 2021.

[5] Y.-S. Yang et al., "A physical-based artificial neural networks compact modeling framework for emerging fets," *IEEE Transactions on Electron Devices*, vol. 71, no. 1, pp. 223–230, 2023.

[6] G. Choe et al., "Machine learning-assisted statistical variation analysis of ferroelectric transistor: From experimental metrology to adaptive modeling," *IEEE Transactions on Electron Devices*, vol. 70, no. 4, pp. 2015–2020, 2023.

[7] T. Mikolajick et al., "The past, the present, and the future of ferroelectric memories," *IEEE Transactions on Electron Devices*, vol. 67, no. 4, pp. 1434–1443, 2020.

[8] J. Wu et al., "Deep learning of accurate force field of ferroelectric hfo 2," *Physical Review B*, vol. 103, no. 2, p. 024 108, 2021.

[9] H. Mulaosmanovic et al., "Ferroelectric field-effect transistors based on HfO $_2$: A review," en, *Nanotechnology*, vol. 32, no. 50, p. 502 002, Dec. 2021.

[10] T. Mikolajick et al., "Feram technology for high density applications," *Microelectronics Reliability*, vol. 41, no. 7, pp. 947–950, 2001.

[11] M. Abuwasib et al., "Cmos compatible integrated ferroelectric tunnel junctions (ftj)," in *2015 73rd Annual Device Research Conference (DRC)*, IEEE, 2015, pp. 45–46.

[12] S. Beyer et al., "FeFET: A versatile CMOS compatible device with game-changing potential," en, in *2020 IEEE International Memory Workshop (IMW)*, Dresden, Germany: IEEE, May 2020, pp. 1–4.

[13] G. Feng et al., "A ferroelectric fin diode for robust non-volatile memory," *Nature Communications*, vol. 15, no. 1, p. 513, 2024.

[14] A. I. Khan et al., "Negative capacitance in a ferroelectric capacitor," *Nature materials*, vol. 14, no. 2, pp. 182–186, 2015.

[15] P. Jiang et al., "Wake-up effect in hfo2-based ferroelectric films," *Advanced Electronic Materials*, vol. 7, no. 1, p. 2 000 728, 2021.

[16] A. Chouprik et al., "Wake-Up in a Hf $_{0.5}$ Zr $_{0.5}$ O $_2$ Film: A Cycle-by-Cycle Emergence of the Remnant Polarization via the Domain Depinning and the Vanishing of the Anomalous Polarization Switching," en, *ACS Applied Electronic Materials*, vol. 1, no. 3, pp. 275–287, Mar. 2019.

[17] R. Y. Choi et al., "Introduction to machine learning, neural networks, and deep learning," *Translational Vision Science & Technology*, vol. 9, no. 2, p. 14, 2020.

[18] T. Hastie et al., "The elements of statistical learning: Data mining, inference, and prediction," *Springer Science and Business Media*, 2009.

[19] I. H. Sarker, "Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions," eng, *SN computer science*, vol. 2, no. 6, p. 420, 2021.

[20] N. Bellarmino et al., "A Multilabel Active Learning Framework for Microcontroller Performance Screening," en, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 42, no. 10, pp. 3436–3449, Oct. 2023.

[21] N. Bellarmino et al., "Deep learning strategies for labeling and accuracy optimization in microcontroller performance screening," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, pp. 1–1, 2024.

[22] M. Long et al., "Learning transferable features with deep adaptation networks," in *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, 2015, pp. 97–105.

[23] F. Zhuang et al., "A Comprehensive Survey on Transfer Learning," *Computing Research Repository (CoRR)*, vol. abs/1911.02685, 2019. arXiv: 1911.02685.

[24] Q. Y. Chen et al., "Artificial neural network compact model for tfts," *2016 7th International Conference on Computer Aided Design for Thin-Film Transistor Technologies (CAD-TFT)*, pp. 1–1, 2016.

[25] M. Li et al., "Physics-inspired neural networks for efficient device compact modeling," *IEEE Journal on Exploratory Solid-State Computational Devices and Circuits*, vol. 2, pp. 44–49, 2016.

[26] J. Wang et al., "Artificial neural network-based compact modeling methodology for advanced transistors," *IEEE Transactions on Electron Devices*, vol. 68, pp. 1318–1325, 2021.

[27] C.-T. Tung et al., "Neural network-based and modeling with high accuracy and potential model speed," *IEEE Transactions on Electron Devices*, vol. 69, no. 11, pp. 6476–6479, 2022.

[28] Y.-S. Yang et al., "A physical-based artificial neural networks compact modeling framework for emerging fets," *IEEE Transactions on Electron Devices*, vol. 71, no. 1, pp. 223–230, 2024.

[29] Microsoft, *Neural Network Intelligence*, version 2.0, Jan. 2021.

[30] T. O. Hodson, "Root mean square error (rmse) or mean absolute error (mae): When to use them or not," *Geoscientific Model Development Discussions*, vol. 2022, pp. 1–10, 2022.

[31] A. F. Agarap, "Deep learning using rectified linear units (relu)," *arXiv preprint arXiv:1803.08375*, 2018.

[32] D. P. Kingma, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[33] R. Hecht-Nielsen, "Theory of the backpropagation neural network," in *Neural networks for perception*, Elsevier, 1992, pp. 65–93.

[34] L. Prechelt, "Early stopping-but when?" In *Neural Networks: Tricks of the trade*, Springer, 2002, pp. 55–69.

[35] C.-T. Tung et al., "A compact model of polycrystalline ferroelectric capacitor," *IEEE Transactions on Electron Devices*, vol. 68, no. 10, pp. 5311–5314, 2021.

[36] Z. Cheng et al., "Neural network approach for ferroelectric hafnium oxide phase identification at the atomistic scale," *Materials Today Electronics*, vol. 3, p. 100 027, 2023.

[37] T. Li et al., "Artificial neural network models for metal-ferroelectric-insulator-semiconductor ferroelectric tunnel junction memristor," *Microelectronics Journal*, vol. 144, p. 106 083, 2024.

[38] J. He et al., "Machine learning identified materials descriptors for ferroelectricity," *Acta Materialia*, vol. 209, p. 116 815, 2021.

[39] T. Wu et al., "Speed up quantum transport device simulation on ferroelectric tunnel junction with machine learning methods," *IEEE Transactions on Electron Devices*, vol. 67, no. 11, pp. 5229–5235, 2020.

[40] K. Ni et al., "A Circuit Compatible Accurate Compact Model for Ferroelectric-FETs," in *2018 IEEE Symposium on VLSI Technology*, ISSN: 2158-9682, Jun. 2018, pp. 131–132.

[41] K. Ni et al., "On the Channel Percolation in Ferroelectric FET Towards Proper Analog States Engineering," en, in *2021 IEEE International Electron Devices Meeting (IEDM)*, San Francisco, CA, USA: IEEE, Dec. 2021, pp. 15.3.1–15.3.4.

[42] C. Alessandri et al., "Monte Carlo Simulation of Switching Dynamics in Polycrystalline Ferroelectric Capacitors," en, *IEEE Transactions on Electron Devices*, vol. 66, no. 8, pp. 3527–3534, Aug. 2019.

[43] D. Zhou et al., "Wake-up effects in si-doped hafnium oxide ferroelectric thin films," *Applied Physics Letters*, vol. 103, no. 19, 2013.

[44] E. S. Muckley et al., "Interpretable models for extrapolation in scientific machine learning," *Digital Discovery*, vol. 2, no. 5, pp. 1425–1435, 2023.

[45] M. W. Browne, "Cross-validation methods," *Journal of mathematical psychology*, vol. 44, no. 1, pp. 108–132, 2000.

[46] T. Cui et al., "Can interface layer be really free for hf x zr 1-x o 2 based ferroelectric field-effect transistors with oxide semiconductor channel?" *IEEE Electron Device Letters*, 2024.