An AI-Driven Framework for Enhancing Resilience in Propagation Models to Enable Digital Twin

Waseem Raza*, Fahd Ahmed Khan*, Haneya Naeem Qureshi*, Usama Masood* and Ali Imran*‡
‡James Watt School of Engineering, University of Glasgow, United Kingdom.
*AI4Networks Research Center, School of Electrical & Computer Engineering, University of Oklahoma, OK, USA
Email: {waseem, fahd.khan, haneya, usama.masood, ali.imran}@ou.edu

***Abstract*—The evolution of wireless cellular networks to support Digital Twins (DTs) requires robust propagation models. Traditional propagation modeling methods, though fundamental, lack the realism, completeness, and computational efficiency required for effective DT synchronization. This inadequacy underscores the need for models that can seamlessly integrate with real-world network dynamics. Therefore, this work critically examines the resilience of conventional machine learning based models and highlights their vulnerabilities to data scarcity and the inherent dynamism of wireless networks. To address these challenges, we propose an innovative approach that leverages a multi-stage GAN for the generation and augmentation of tabular synthetic data, coupled with an Attention through Segmentation training strategy. This strategy is based on partitioning the data distribution based on the histogram of important features and replacing the single complex model with multiple simpler models focused on specific parts of the histogram. This dual approach significantly improves the resilience of the model, and our evaluations in realistic scenarios show an impressive recovery of more than 90% performance loss compared to traditional models and this improvement is achieved with a notable reduction in model complexity. Our research marks a significant advancement in the development of resilient and efficient propagation models for the next generation of wireless networks.**

***Index Terms*—Propagation Modeling, Digital Twins, Resilience Analysis, Data Augmentation, Attention through Segmentation.**

I. Introduction

Digital Twin (DT) is envisioned to play a pivotal role in the advancement of emerging cellular networks. A DT is a software replica of the mobile network that facilitates continuous prototyping, testing, and optimization [1], [2]. For example, DT can be used to analyze performance and assess the impact of new services on existing ones [1]. These are useful for two main purposes: (1) teaching AI systems to handle situations where real networks have limited available data for experimentation and solution design, and (2) evaluating and improving new AI solutions before their implementation in real networks, which is otherwise a more risky undertaking [1], [2]. A key element in the advancement of DT creation involves modeling the propagation of the wireless network, which must embody key characteristics such as realism, computational efficiency, and seamless integration with the physical wireless network of the real world [2].

Traditional propagation models, including empirical, deterministic, and stochastic, fail to comply with the aforementioned key characteristics of DTs. For example, the realism in propagation models is realized through deterministic simulations based on ray-tracing [3]. However, such simulations are highly computationally inefficient, especially when taking into account the temporal dynamic nature of wireless systems. This makes the integration of such models with real networks infeasible. On the contrary, empirical and stochastic approaches such as COST-Hata and ITU-R P.453-15 [4], although computationally efficient, deviate from the realistic physical and geographic structure of the environment [3]. In response to the limitations of traditional propagation models, data-driven Machine Learning (ML) techniques and Deep Neural Networks (DNNs), which make use of large-scale datasets, are emerging as promising solutions [5], [6].

A. Related Work

The exploration of ML in wireless network propagation modeling and pathloss prediction research has covered multiple environments and methodologies [7]–[10]. In particular, [7] validates the use of DNNs for the prediction of pathloss in macrocells and different terrains, while [8] improves the design of ANN through a composite differential evolution algorithm, leading to better prediction accuracy. The integration of environmental factors into pathloss models using learning machines and DNN is detailed in [9], highlighting the importance of such features for heterogeneous networks. Furthermore, [10] emphasizes not only accurate prediction of network coverage but also model interpretability, which is crucial for practical applications. Advances in 3D propagation modeling by [3] tackle complex spatial dynamics and combine AI interpretability with comprehensive 3D modeling, proposing sophisticated tools for autonomous network planning.

Despite these advances, machine learning-based propagation models face significant challenges, mainly due to the scarcity of diverse large-scale datasets [11]–[13]. This scarcity, coupled with data imbalance and representation issues, can cause the models to inclined toward a particular propagation environments, such as urban propagation patterns, undermining their effectiveness in rural or varied terrains. This phenomenon, known as the *distribution shift*, poses a substantial challenge to the generalizability of the model [14], which requires the development of models that can adapt and perform in a wide spectrum of environments [11].

The concept of distribution shift is studied across different data types and generalization strategies [15], [16]. For example, advances in meta-learning, methods based on Generative Adversarial Networks (GANs), and ensemble learning offer promising avenues for improving model resilience to distribution shifts [17]. However, these strategies often overlook the unique challenges of wireless networks, such as the intricacies of channel dynamics and operational parameters. The current literature suggests a research gap in data-driven propagation models that are resilient to realistic scenarios and distribution changes, tailored for emerging wireless networks [17]. This gap underscores the need for dedicated efforts to develop ML-based models that effectively address the nuances of wireless network propagation, ensuring robust performance across diverse conditions and environments.

B. Contributions

To address the gaps mentioned above, this study aims to address data-driven propagation modeling for 5G and beyond networks and proposes a robust and resilient ML model training framework for real-world network environments that have distribution shifts, unrepresentative data, and data scarcity. The key contributing points are summarized here.

- We propose a noval data-driven propagation modeling framework to effectively predict the Reference Signal Received Power (RSRP) with respect to network environment features, even when facing practical scenarios with test data distributions vastly different from that of training data, ensuring resilience against unseen scenarios.
- This model resilience-enhancing approach is twofold, first using a Conditional Tabular GAN (CTGAN) to augment the training data and get a balanced distribution of key features. This step provides foundational robustness, preparing the model to deal with real-world Non-Identical Distributions (NID), and helps to prevent mode collapse. Then, we apply an Attention through Segmentation (AtS) strategy by training simpler models for distinct data segments, thereby improving focus on specific data ranges and enhancing the model's capability to identify localized patterns and increase model accuracy.
- The framework promotes rapid convergence by integrating a Multi-stage GAN (MGAN) that benefits from weight transfer from a pre-trained DNN to a conditional GAN generator. This strategy accelerates the optimal weight-finding process during training, offering an advantage over traditional GANs, which start with randomly initialized generator weights.
- We rigorously evaluated the quality of the synthetic data produced by MGAN versus conventional GANs using various metrics. Furthermore, we compare the performance of our AtS method with baseline models in different realistic test scenarios, exploring the balance between the performance of the Root Mean Square Error (RMSE), the complexity of the model, and the convergence time. Our findings indicate a significant recovery from performance loss, over 90%, and a reduction in the GAN convergence time of approximately 50%, compared to traditional methods.

II. Proposed Framework Overview

The proposed framework comprises five interconnected modules, as illustrated in Fig. 1. In *Data Collection* module (M1), model training and test data are acquired from real operator networks or network simulations. This data is fed into the *Model Training* module (M2), which focuses on training various Machine Learning (ML) models, along with preprocessing tasks, feature scaling, and optimizing hyperparameters optimization. The trained models then undergo resilience testing in the *Model Resilience Testing* module (M3) to validate performance against both ID and NID test scenarios. The performance of a model is considered unsatisfactory if the RMSE performance differs significantly, $> 25\%$ in our analysis, between the ID and NID tests.

In the cases of poor performance of the model, the framework moves to the module *New Data Generation and Augmentation* (M4) module for data augmentation, employing GANs for synthetic data generation and integration with original data. The final module, *Attention through Segmentation* (M5), focuses on data-dependent model selection, facilitating the training of simpler attention models in specific bins of the histogram. One of its blocks stores the trained models along with the basic data statistics (range and mean) used for training. During inference, these statistics help map the test data to the most appropriate models for predictions. The intricacies of each module are explored in detail in the following sections.

III. Resiliency Analysis of Existing Models

Before delving into the details of model resilience analysis carried out in this work, we discuss the data collection and feature engineering to get the right data used for training these model under consideration for resilience analysis.

A. Data Collection and Feature Engineering: (M1 in Fig. 1)

A realistic commercial planning tool based on ray tracing is used to create a sophisticated network topology and generate multiple raw data sets [3]. The network topology incorporates actual antenna heights, geographic data (ground heights, building heights and land use maps), and the Aster propagation model [18]. The raw data set comprises data relevant to signal propagation, collected from the commercial planning tool, and classified into three main groups: (1) site-specific information (including location and antenna details), (2) geographic information (encompassing terrain, building, and land cover data), and (3) user measurements (comprising RSSI, location, and network information)[1]. Feature engineering is critical to improve the performance of machine learning models. Therefore, in our study, using domain knowledge, we transform raw site-specific, geographic and user data from BS into relevant characteristics, including distances, clutter information, penetrations of buildings, diffraction points and angular separations between BSs and users[1].

[1]For more details about these datasets refer to Section II-C of [3].

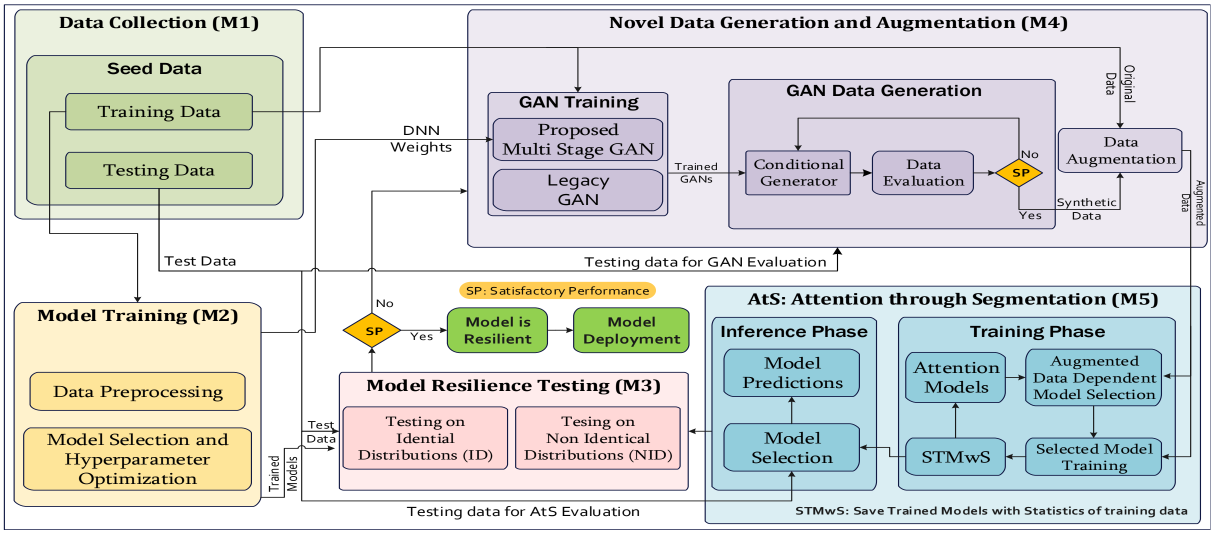

Fig. 1: The proposed framework for the resilience analysis of data-driven propagation models (M1-to-M3), and resilience enhancement of these models using CTGAN based data augmentation (M4) and Attention through Segmentation (M5).

B. *Model Resilience Analysis (M1-to-M3 in Fig. 1)*

To test the resilience of the model, we use sensitivity analysis to select the high impact features and their distributions to select the NID test cases.

1) SHAP-based High Impact Features' Distributions: The Shapley Additive Explanations (SHAP) concept is used to quantify the contribution of individual characteristics to the prediction of ML models. We applied the LightGBM model, a prevalent choice in propagation model research, to perform the SHAP analysis of the features. This analysis identifies three main features, given in descending order of importance here: *Distance (F1)*, *Indoor Path (F2)*, and *BS-Azimuth (F3)*. Our findings indicate that the F1 feature exhibits a positive skew, with a notable concentration of user equipment approximately 500 meters from the base stations. The Indoor Path feature follows an exponential distribution, predominantly representing shorter distances and suggesting that a majority of signals traverse minimal indoor distances or obstructions, with zero values indicating direct line-of-sight conditions. The BS Azimuth exhibits a bimodal distribution that highlights the variability in BS orientations. To analyze the resilience of the model, we discretize these feature distributions into histograms with bin sizes of B_S (6, 12 bins), defining NID scenarios using test data from the extremities of the histogram.

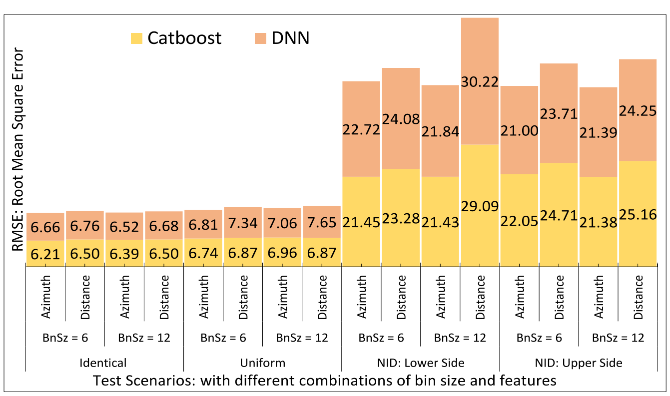

Fig. 2: RMSE performance of ML models for identical, uniform, and two realistic non-identical test scenarios, where test data is sampled from the upper and lower sides of the distribution of relevant features.

2) Model Testing against Various Test Scenarios (M3): Leveraging SHAP-identified features, we evaluate the resilience of the model in realistic scenarios, focusing on consistency of performance in changes in data distribution. For this purpose, we use the Catboost and DNN models to evaluate performance in various test scenarios, with results depicted in Fig. 2. Four different test scenarios are examined; Identical (ID), uniform (equal representation from different bins of the histogram), and NID (Lower Side and Upper Side) scenarios featuring test data exclusively from the histogram extremities. Initial findings indicate a comparable performance between DNN and Catboost in ID and uniform cases, as depicted in Fig. 2. However, performance notably declines under NID conditions, underscoring a resilience gap in current ML models and the need for improved model resilience strategies.

IV. CTGAN based Data Augmentation (M4) and Attention through Segmentation (M5)

A. *Novel Data Generation Approach (M4 in Fig. 1)*

The M4 module in Fig. 1 aims to create synthetic tabular data, a task more challenging than typical image augmentation through GANs, due to various types of feature and distributions of tabular data features and correlations between them. CTGAN is a specialized GAN architecture engineered to address the complexities of generating tabular data. For our problem of data generation for distribution balancing, we propose improvements in CTGAN, discussed in the following.

1) Proposed Improvements in CTGANs: Our data generation approach includes two improvements in the CTGAN training and data generation phases. First, the CTGAN is initialized with the weights from a pre-trained DNN model in M2. Doing so speeds up the convergence of CTGAN and helps avoid modal collapse, without compromising the quality of generated data. Then, we propose an iterative approach to condition all available values in a particular bin of the histogram, and a mathematical formulation to manage the count of each conditioned value, processed in the *Conditional Generation* block of the data generation phase in M4. This also involves the *Data Evaluation* block, which evaluates the quality of the generated data by comparing it with the respective test data, using various performance evaluation metrics [19]. Synthetic data is enhanced with original data only if it has *Satisfactory Performance* (SP), measured by applying a threshold to GAN evaluation metrics, such as the quality score and the coverage report [19], see Section V-B.

2) Mathematics of Feature Distribution Balancing: The number of values that need to be augmented to make a histogram bin similar in size to other bins depends on various factors, such as the underlying distribution of features, the number of bins in the histogram and the overall size of the data. If we have a dataset of size D_l values and due to the underlying distribution, its most important features can have some bins with sufficiently high values count, while other bins with very few values. For a given bin size B_s, we define the *balance threshold* $B_t = D_l/B_s$, as the minimum value count for each bin so that it can be considered balanced. For a given bin b_i, where $i \in Z_s$ where $Z_s = \{x \in \mathbb{Z} \mid 1 \leq x \leq B_s\}$, the count of a particular feature F in b_i is shown by V_i, and the number of all bins with $V_i < B_t$ termed Low-frequency Bins (*LfB*) is shown by N_b^l. Therefore, depending on the difference in N_b^l and its individual values $|B_t - V_j|$, where $j \in Z_l$ and $Z_l = \{x \in \mathbb{Z} \mid 1 \leq x \leq N_b^l\}$, we determine the total count of additional values, $N_g^v = \sum_{k=1}^{N_b^l} |B_t - V_k|$ required to make LfB more balanced and representative.

3) CTGAN Implementation in Synthetic Data Vault (SDV): We leverage the CTGAN implementation from the SDV library, a versatile GAN capable of handling mixed data types and imbalanced datasets [20]. CTGAN employs a conditional sampling approach, enabling data generation post-training without predefined conditions. It aims to accurately model the data distribution, grasping intricate inter-feature relationships and fostering a model adaptable to conditions applied during sampling. This approach offers several advantages, for example, generating diverse datasets from a single model and providing a cost-effective solution to manage data variances.

4) CTGAN Data Balancing Capabilities and Limitations: Despite CTGAN's capabilities, it cannot fully balance the histogram bins, especially those that contain very lower values for conditioning. The upper half of Table I shows the number of LfB in the original and augmented data for three features. We have defined a parameter, PoLfB, which stands for the 'Proportion' of values present in LfB to the overall data size, and data augmentation impact is represented by the change in PoLfB and NoLfB values as we transition from original to augmented data rows for a given bin size and feature. Although CTGAN effectively increases PoLfB to a maximum value of 0.999 for F1 and F3, for F2, it reaches 0.840. This is because some bins of the F2 histogram have very low values and the GAN conditioning is limited by the values in a particular bin.

B. AtS: Attention through Segmentation (M5 in Fig. 1)

Now we move on to the proposed AtS approach, focusing on its motivational and foundational aspects, and implementation of the AtS module (M5) in the proposed framework in Fig. 1.

1) Motivation and Implementation of AtS: We employ the AtS scheme to improve the resilience and precision of the model, training multiple specialized models on different data segments instead of a single complex model [21]. This method improves the efficiency and precision of the regression model training by focusing on specific value ranges, ensuring robustness and reducing computational load. In the AtS module (**M5**), models are trained according to the width of the histogram bin, with k ML models covering each of the bins N/k, where k is a factor of N. This approach ranges from a conventional single model ($k = 1$) to one model per bin ($k = N$), allowing for adjustable attention levels.

2) AtS Module in Proposed Framework: The **M5** module's *Augmented Data Dependent Model Selection* block plays a crucial role in identifying the optimal model for specific augmented data segments, considering bin size, attention level, and feature distribution. Concurrently, the *Attention Models* block is tasked with devising the architecture for various attention models as per the Model Selection block's directives. Once selected, a model undergoes training for its designated data segment, with the *STMwS* (Save Trained Model with Statistics) storing the model along with its training data statistics. In the inference phase, these pre-trained models and their statistics in STMwS are leveraged for RMSE prediction on new test data, ensuring the selection of suitable models through a comparison of stored and test data statistics.

Table I: Comparison of GAN data generation and convergence: NoLfB (PoLfB): Number (Percentage) of Low-frequency Bins, GCI: GANs Convergence Iterations, DGT: Data Generation Time.

		F1: Distance	F2: Indoor Path	F3: BS Azimuth
	Data Type	**NoLfB/PoLfB**	**NoLfB/PoLfB**	**NoLfB/PoLfB**
GAN Data Generation	**Orig. Data**	3/0.745	4/0.495	2/0.841
	Aug. Data	0/0.999	1/0.840	0/0.999
	GAN Type	**GCI/DGT**	**GCI/DGT**	**GCI/DGT**
GAN Convergence	**LGAN**	289/1615	323/1870	183/1254
	MGAN	188/1574	201/1988	128/1199

V. Simulation Setup and Performance Evaluation

In this section, we discuss the experimental setup, evaluation of GAN-based data generation, and proposed AtS based approach to improve the resilience of propagation models.

A. Experimental Setup

Using the Atoll commercial planning tool [18], based on 3D ray tracing, we simulate a 3.8 sq. Km network in central Brussels, Belgium, with 10 macrocell base stations, detailed in Table II. The area is divided into bins to calculate the average

Table II: Network simulation parameters with their values.

Parameters	Values
Network layout	10 macrocell sites
Number of users in simulation area	10000 users
Path loss model	Aster (ray tracing)
BS maximum transmit power	43dBm
BS antenna gain	18.3 dBi
Channel bandwidth	5 MHz
Size of the simulated data	300 K
Training and test data size	150 K
Generator hidden layers shapes	[64, 64, 64, 64, 64]
Discriminator hidden layers shapes	[64, 64, 64, 64, 64]

RSRP values per bin of 10K Poisson distributed users. This data set includes 9 features, 1 target variable (RSRP), and 300K instances, with 150K reserved for testing. Different sizes of test data are derived from this set for scenarios M2, M3, and M4. For DNN, we use 120K instances for training, 30K for validation, and sample another 30 K from the test set. The GAN training dataset matches this, but the testing datasets are adjusted to ensure fair evaluations due to the variable generated dataset sizes based on features and bin dimensions.

B. GAN Data Generation Performance Evaluation

The effective use of GANs depends on evaluating their data generation performance, which we assess using two key metrics: the ‘quality score’ and the ‘diagnostic report’ [19]. The quality score focuses on the precision of the generated data by examining the shapes of individual columns and the pairwise trends, while the diagnostic report provides a broader evaluation, considering the coverage, synthesis, and limits of each column (see [19] for more details). Both GANs exhibit comparable quality scores, the proposed MGAN outperforms the LGAN in quality score and in the synthesis of unique values, as shown in Fig. 3. Although the coverage and boundary metrics are slightly better for LGAN, they are significantly higher and comparable for MGAN, suggesting that both GANs can generate high-quality data.

Analyzing convergence performance (see Table I), MGAN shows superiority over LGAN in the GAN Convergence Iteration (GCI) and data generation time (DGT), benefiting from modified training and initial weight transfer. The significant improvement in GCI and DGT, especially for features like F2: Indoor Path, is indicative of the GAN’s ability to bridge the gap in data augmentation, with the gap reduction in PoLfB quantified as $0.840 - 0.495 = 0.345$ for F2 in the upper half of Table I. Trends in these metrics, along with performance in other features such as F1 and F3, provide information on the effectiveness of GAN-based approaches in improving data representation and without compromising on its quality.

C. Evaluation of Proposed AtS Methodology

The performance of the proposed AtS schemes with varying depth of attention levels is shown in Fig. 4. Here, A01M is similar to conventional training, where no segments are made and only one model is trained for the entire data range; hence it becomes our *baseline*. Similarly, for $B_s = 6$, A02M has 2 models that span 3 bins each and A03M has 3 models that span 2 bins each. Fig. 4, a stacked grouped bar graph, shows all

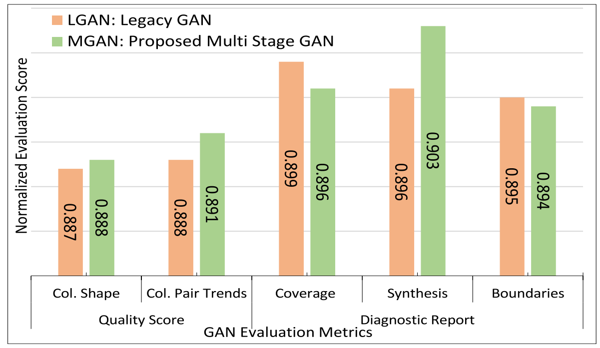

Fig. 3: GAN quality comparison for both LGAN and MGAN.

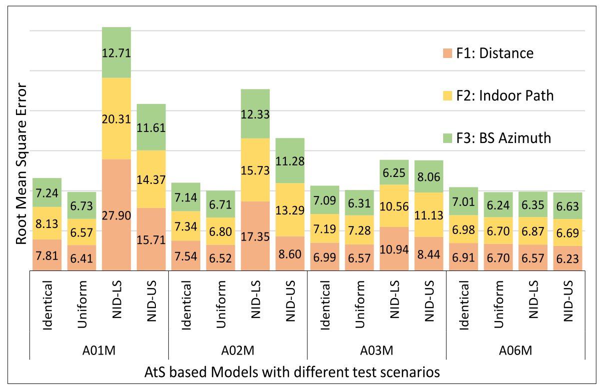

Fig. 4: Testing the RMSE performance for baseline (A01M) and AtS models (A02M, A03M, A06M) for $B_S = 06$.

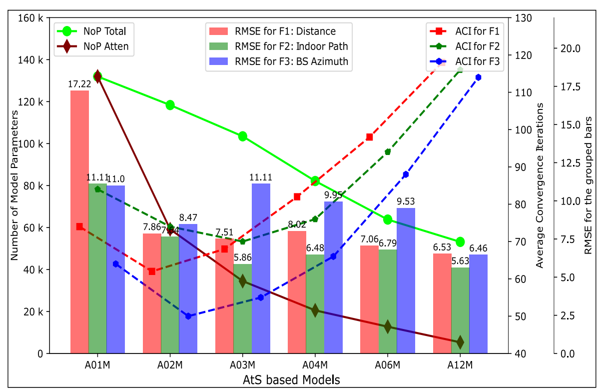

Fig. 5: Trade-off between model complexity shown as number of parameters (NoP), improved resiliency (RMSE improvements), and convergence performance for $B_S = 12$.

possible levels of attention for $B_S = 6$, each attention level showing 4 different test cases, and each test case shows a stacked bar representing the RMSE values for three important characteristics under consideration.

1) DNN Models RMSE Comparison on Augmented Data: An analysis of all features reveals that the proposed AtS approach is effective, as performance (or RMSE) improves as the models become more specialized, from A01M to A06M. This is especially evident in NIDs such as “LS: lower side”, where the RMSE decreases for all three features, for example, from 27.90 in A01M to 6.57 in A06M for *Distance* feature, indicating the potential robustness of AtS-based specialized models. The consistent decrease in RMSE in all test scenarios suggests the potential benefit of the AtS-based localized specialized training approach. When looking at the individual features, each has its nuances and trends. For example, for a given NID test scenario, the *Distance* feature shows a steady and consistent decrease in RMSE in all attention cases and test scenarios, suggesting a consistent advantage of AtS-based training. However, when considering the features of *Indoor Path* and *BS Azimuth* features, the rate of reduction of RMSE from A01M to A06M is more abrupt in a certain NID test case. This variation in RMSE for different features is mainly related to their distribution variety and their importance order in model prediction.

2) Model Parameters and Performance Trade-Offs: This analysis explores the correlation between the complexity of the AtS model (manifested as NoP: Number of Parameters), Average Convergence Iterations (ACI), and RMSE (Fig. 5). Focusing on a bin size $B_S = 12$ and a consistent size of the DNN model, the study finds that specialized attention schemes reduce NoP and improve RMSE performance, indicating efficient model learning with lower complexity. The best performance is achieved with A12M, using only parameters 5316 per DNN model, totaling about $60K$ parameters, significantly less than the baseline of $130K$. Although AtS depth/level decreases ACI, suggesting faster learning, a trend shift occurs where higher attention levels lengthen the DNN model convergence. A12M shows ACI values approximately twice those of the baseline, but improves resilience without overburdening training. The suitability of AtS for parallel training effectively addresses these higher ACI values.

VI. Conclusion and Future work

In this article, we quantify and address key issues in current propagation modeling to enable digital twins, that is, lack of resilience caused by distribution shift and data scarcity issues in wireless networks. We achieve resilience against realistic test-case scenarios by proposing a framework that uses GAN-based data augmentation to balance the distribution of influential features and attention through segmentation-enhanced training to address the performance degeneration issue. We also introduce a novel multistage GAN module based on transfer learning to reduce convergence time. Our approach, compared to the baseline scheme for different realistic test cases, shows recovery of more than 90% performance loss compared to its conventional counterpart, with the additional benefit of reduced model complexity. In the future, our aim is to extend this work to different data sizes and histogram bins.

VII. Acknowledgment

This work is supported by the National Science Foundation through grant number 1923669.

References

[1] X. Li, Y. Zhang, W. Zhang, Y. Zhang, X. Wang, and Y. Wang, "Digital-Twin-Enabled 6G: Vision and Requirements," *arXiv preprint arXiv:2102.12169*, 2021.

[2] F. Lara, R. Lara-Cueva, M. Castillo, J. F. Arellano, and L. Top'on, "Modeling Wireless Propagation Channel: A Traditional Versus Machine Learning Approach," in *Emerging Research in Intelligent Systems*. Springer, 2022, p. 281–297.

[3] U. Masood, H. Farooq, A. Imran, and A. Abu-Dayya, "Interpretable AI-Based Large-Scale 3D Pathloss Prediction Model for Enabling Emerging Self-Driving Networks," *IEEE Transactions on Mobile Computing*, vol. 22, no. 7, pp. 3967–3984, 2023.

[4] G. R. Maurya, P. A. Kokate, S. K. Lokhande, and J. A. Shrawankar, "Review on Investigation and Assessment of Path Loss Models in Urban and Rural Environment," *IOP Conference Series: Materials Science and Engineering*, vol. 225, p. 012219, 2017.

[5] D. F. S. Fernandes, A. Raimundo, F. Cercas, P. J. A. Sebasti ao, R. Dinis, and L. S. Ferreira, "Comparison of Artificial Intelligence and Semi-Empirical Methodologies for Estimation of Coverage in Mobile Networks," *IEEE Access*, vol. 8, p. 139803–139812, 2020.

[6] S. Sun and S. Mao, "DeepProp: A Deep Learning Framework for Wireless Propagation Modeling Using Ray Tracing Data," *IEEE Transactions on Wireless Communications*, vol. 19, no. 9, p. 6156–6170, 2020.

[7] M. Ribero, R. W. Heath, H. Vikalo, D. Chizhik, and R. A. Valenzuela, "Deep Learning Propagation Models over Irregular Rerrain," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 4519–4523.

[8] S. P. Sotiroudis, S. K. Goudos, K. A. Gotsis, K. Siakavara, and J. N. Sahalos, "Application of a Composite Differential Evolution Algorithm in Optimal Neural Network Design for Propagation Path-Loss Prediction in Mobile Communication Systems," *IEEE Antennas and Wireless Propagation Letters*, vol. 12, pp. 364–367, 2013.

[9] S. I. Popoola, E. Adetiba, A. A. Atayero, N. Faruk, and C. T. Calafate, "Optimal Model for Path Loss Predictions using Feed-Forward Neural Networks," *Cogent Engineering*, vol. 5, no. 1, p. 1444345, 2018.

[10] A. Ghasemi, "Data-driven prediction of cellular networks coverage: An interpretable machine-learning model," in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2018, pp. 604–608.

[11] V. P. Rekkas, S. Sotiroudis, P. Sarigiannidis, S. Wan, G. K. Karagiannidis, and S. K. Goudos, "Machine Learning in Beyond 5G/6G Networks—State-of-the-Art and Future Trends," *Electronics*, vol. 10, p. 2786, 2021.

[12] M. Chen, Y.-C. Liang, and H.-H. Chen, "Machine Learning for Wireless Networks with Artificial Intelligence: A Comprehensive Survey," *IEEE Access*, vol. 6, pp. 36 114–36 134, 2018.

[13] H. N. Qureshi, U. Masood, M. Manalastas, S. M. A. Zaidi, H. Farooq, J. Forgeat, M. Bouton, S. Bothe, P. Karlsson, A. Rizwan *et al.*, "Towards Addressing Training Data Scarcity Challenge in Emerging Radio Access Networks: A Survey and Framework," *IEEE Communications Surveys & Tutorials*, 2023.

[14] M. Zhang, "Adaptation Based Approaches to Distribution Shift Problems," Ph.D. dissertation, EECS Department, University of California, Berkeley, Dec 2021. [Online]. Available: http://www2.eecs.berkeley.edu/Pubs/TechRpts/2021/EECS-2021-262.html

[15] "awesome-distribution-shift," accessed: February 29, 2024. [Online]. Available: https://github.com/weitianxin/awesome-distribution-shift#learning-strategy

[16] S.-W. Huang, C.-T. Lin, S.-P. Chen, Y.-Y. Wu, P.-H. Hsu, and S.-H. Lai, "AugGAN: Cross Domain Adaptation with GAN-Based Data Augmentation," in *European Conference on Computer Vision*. Springer, 2018, pp. 731–744.

[17] M. Akrout, A. Feriani, F. Bellili, A. Mezghani, and E. Hossain, "Domain Generalization in Machine Learning Models for Wireless Communications: Concepts, State-of-the-Art, and Open Issues," *arXiv preprint arXiv:2303.08106*, 2023.

[18] "Atoll," [Online]. Available: https://www.forsk.com/atoll-overview.

[19] "SDMetrics: Synthetic Data Metrics," accessed: February 29, 2024. [Online]. Available: https://docs.sdv.dev/sdmetrics/

[20] "Synthetic Data Vault," accessed: February 29, 2024. [Online]. Available: https://docs.sdv.dev/sdv/

[21] C. Zhang, S. Dang, B. Shihada, and M.-S. Alouini, "Dual Attention-Based Federated Learning for Wireless Traffic Prediction," in *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, 2021, pp. 1–10.