AI for Link Adaptation and Energy Prediction in Realistic 5G Networks

Alka Valiparambil Narendran

University of Lorraine, France — Leeds Beckett University, UK — Luleå University of Technology, Sweden

Abstract: As 5G networks grow in complexity, ensuring efficient spectrum use and energy consumption has become a critical challenge—one where AI offers significant promise. This paper explores AI-driven optimization of link adaptation in 5G standalone networks, focusing on MCS prediction using SINR and CQI as input. With more than 20,000 samples generated through 5G-LENA simulations, several supervised models, including CNN, LSTM, and contextual variants, were evaluated. While overall classification accuracy remained modest (52–55%), recurrent models outperformed others, aligning with prior research on temporal dependencies in wireless channels.

Parallel efforts in energy modeling used a public dataset of 5G base station activity. Traditional regressors struggled with high error, but a neural model inspired by Chen and et al. (2024). achieved a MAPE of 5.5%, underscoring the importance of temporal and contextual features in forecasting energy use.

Together, these results highlight the practical potential of AI in 5G network optimization while emphasizing the need for richer datasets and more temporally expressive models.

Keywords: 5G Mobile Communication, Link Adaptation, Modulation and Coding Scheme (MCS), Artificial Intelligence, Machine Learning, Energy Efficiency, Radio Access Networks (RAN), Time-Series Prediction, Wireless Systems Modeling, Sustainable Communication Systems

1. INTRODUCTION

In today's rapidly evolving digital landscape, the shift to fifth-generation (5G) mobile networks has brought dramatic improvements in speed, latency, and connectivity. As this transition accelerates and early explorations of 6G begin, optimizing the performance and energy efficiency of existing 5G infrastructure becomes increasingly critical.

A central mechanism in 5G performance is link adaptation, where transmission parameters adapt to channel conditions. The key decision in this process is the selection of the Modulation and Coding Scheme (MCS). It directly affects throughput, reliability, and energy use. Inefficient MCS selection leads to resource underutilization or unnecessary retransmissions, both of which degrade performance.

This work investigates how artificial intelligence can enhance link adaptation by predicting optimal MCS values based on recent network conditions. Using data generated from the 5G-LENA simulator, AI models were trained to make real-time MCS predictions, focusing on downlink communication, where such decisions are centrally made by the gNB.

To complement this, the study also models energy consumption using a real-world dataset, aiming to predict power usage under varying radio configurations. While developed independently, the two components offer complementary insights into how intelligent systems can balance performance with sustainability.

By addressing both throughput and energy use, this work contributes toward smarter, greener 5G networks, supporting broader goals in sustainable infrastructure and responsible innovation. These are the main research objectives:

RO1: Identify and extract relevant KPIs from 5G-LENA simulations, including SNR, CQI, transmission power, and buffer size, to support AI-based MCS prediction.

RO2: Develop and evaluate an AI model to predict MCS values based on historical network states, targeting improved robustness in the absence of frequent CSI updates.

RO3: Design a separate AI model to estimate energy consumption under varying transmission parameters using an external dataset.

RO4: Analyze and compare energy efficiency patterns derived from the energy model, highlighting trade-offs with performance.

RO5: Integrate findings to formulate practical recommendations for optimizing 5G link adaptation strategies with sustainability considerations.

RO6: Investigate the practical utility of AI-driven optimization techniques for 5G systems, demonstrated within the 5G-LENA simulation framework, and identify pathways for their integration or enhancement in alignment with the objectives of industry partner Effnet.

The remainder of this paper is structured as follows: the next section presents a literature review that situates this work within current research on link adaptation, AI in 5G, and energy efficiency modeling. This is followed by the methodology section, which is divided into two parts—one focusing on AI-based MCS prediction using simulation data, and the other on energy consumption modeling using a real-world dataset. The results and discussion section then evaluates model performance and reflects on the implications of both experiments. Finally, the paper concludes with key insights and outlines directions for future work

2. LITERATURE REVIEW

Artificial Intelligence is rapidly reshaping 5G RAN operations, with applications ranging from traffic forecasting and interference management to energy optimization and autonomous fault detection (Balasubramanian et al. (2021); Tran et al. (2024); Sahu and Sahu (2024)). Techniques like reinforcement learning, federated learning, and hybrid algorithms have shown strong results in dynamic resource allocation and predictive control (Lee et al. (2021); Zhao et al. (2020); Rezazadeh et al. (2023)). These models often power components like the RAN Intelligent Controller (RIC), enabling near real-time optimization through AI-driven loops (Balasubramanian et al. (2021)). However, many of these approaches assume access to complex infrastructure and broad datasets. Given the constraints and the need for measurable outcomes, this study focuses on *link adaptation*, particularly MCS prediction—a Layer 2 task with direct performance impact and practical compatibility with Effnet's ongoing development in 5G scheduler optimization (Abubakar et al. (2020); Lee et al. (2021)).

Link adaptation is essential in 5G networks, dynamically adjusting transmission parameters—such as modulation, coding rate, and transmit power—based on real-time channel conditions. Its aim is to maximize throughput and spectral efficiency while minimizing error rates Gigayasa Wireless (2025). Traditional approaches divide this process into two layers: Inner Loop Link Adaptation (ILLA) and Outer Loop Link Adaptation (OLLA).

ILLA selects the Modulation and Coding Scheme (MCS) based on SINR or CQI values reported by the UE, using lookup tables for rapid decisions in stable environments Gigayasa Wireless (2025). However, this layer often assumes ideal feedback, which may lead to misestimations under rapid channel changes or delayed feedback Sahu and Sahu (2024). OLLA complements this by adjusting a correction offset using ACK/NACK feedback from HARQ transmissions, helping maintain the target BLER of 10% Lee et al. (2021). Yet, OLLA is reactive and may converge too slowly for high-mobility or low-latency applications like URLLC.

Effnet's logs reflect this process: HARQ feedback directly affects MCS via the Link Adaptation Mechanism (LAM), tweaking the MCS up or down. While effective, such rule-based systems lack predictive insight. This motivates the use of AI-based models that learn from SINR, CQI, and HARQ patterns to anticipate channel shifts and optimize MCS decisions more proactively.

AI-driven Modulation and Coding Scheme (MCS) selection generally falls into two main categories: direct and indirect prediction models. Direct approaches map input features such as SINR, CQI, or UE mobility directly to an MCS value, treating the task as a classification problem. Tsipi et al. (2024) evaluated several machine learning methods, finding that an artificial neural network (ANN) with two hidden layers outperformed others in terms of classification accuracy . Similarly, Kojima et al. (2019) proposed a neural network to estimate SNR, which was then used for adaptive MCS adjustment, demonstrating high accuracy even in noisy conditions .

Indirect methods introduce an intermediate step where the model first predicts a channel metric—like SNR or channel impulse response (CIR)—before mapping it to an MCS value using thresholds or lookup tables. Stenhammar et al. (2024) compared several neural network architectures for wireless channel prediction, concluding that recurrent models like GRUs performed better in dynamic scenarios. This approach has the added benefit of interpretability, particularly when forecasting future channel states or reliability under uncertainty.

More recent studies have explored robust MCS selection using temporal deep learning models. Varshney et al. (2023) proposed a deep learning framework based on LSTM networks for predicting CIR in 5G environments, incorporating physical features such as distance and delay spread . Meanwhile, Tsipi et al. (2024) demonstrated a rigorous comparison among ML methods for MCS Prediction, in which LSTM Deep Nueral Network with 2 layers got a 98% accuracy.

Together, these works highlight a shift from static threshold-based systems toward predictive and adaptive models. Whether through direct classification or multi-step channel prediction, AI offers the potential to enhance MCS decision-making under complex and fast-varying radio conditions.

It should also be noted that most of these works focus on traffic in only one direction - either uplink(UL) or downlink(DL). Varshney et al. (2023), Tsipi et al. (2024) and Seeram (2022) focus on DL, while Elgabroun (2019) focus on UL. Since MCS is determined at the gNB side in the downlink, this study also restricts its scope to downlink traffic, enabling a more focused evaluation of channel-driven adaptation strategies.

Table A.1 in Appendix A summarizes the datasets used in recent AI-based link adaptation studies, revealing a reliance on simulation due to the logistical and commercial challenges of collecting live 5G data. However, the level of realism across these simulations varies widely.

Some studies, such as Oh et al. Oh et al. (2023), utilize stepwise SINR inputs to create clean, structured datasets. While effective for training convergence, this approach lacks the volatility and noise typical of real-world deployments. Others aim for higher fidelity. Varshney et al. Varshney et al. (2023), for example, employ NYUSIM—a statistical ray-tracing-based mmWave simulator—to generate spatially aware traces, though it does not fully align with 3GPP standards. Elgabroun Elgabroun (2019) leverages a proprietary Ericsson mmWave simulator, which offers detailed realism but is not publicly accessible for validation or reuse.

The 5G-LENA simulation framework, used in this study, represents a middle ground. Built atop components from NYUSIM and the Madrid Mobility Model (MMB), it integrates 3GPP-compliant features such as HARQ retransmissions, CQI feedback, and realistic urban macrocell propagation. This configuration enables simulations that better reflect the spatio-temporal complexity of real deployments while maintaining transparency and reproducibility—addressing a key limitation in earlier datasets

While the simulation captures realistic link adaptation behavior, it lacks power consumption metrics. To address this, an external dataset of real-world 5G equipment power profiles was used to model energy use across configurations. Unlike static assumptions, this approach reflects variation with frequency, bandwidth, and antenna count. Though not directly integrated with the simulation, it complements the study's dual focus—optimizing both link performance and energy efficiency—aligning with sustainability goals in next-gen RAN design Parsa et al. (2022)Elgabroun (2019).

While energy efficiency has emerged as a critical focus in 5G RAN optimization, many existing studies only integrate it by coupling power control directly with link adaptation mechanisms such as MCS selection Parsa et al. (2022), Norolahi and Azmi (2023). These approaches aim for joint optimization—balancing throughput, power usage, and latency—but often rely on simplified assumptions like continuous CQI feedback or idealized simulations. In contrast, this study adopts a two-pronged methodology: one that models MCS prediction using realistic simulation traces, and another that investigates energy consumption using a separate, empirical dataset. Although these components are not directly integrated, this separation is intentional. It reflects the current limitation in aligning link adaptation with accurate energy profiling, especially in open-source simulation environments like 5G-LENA. Rather than forcing an artificial connection, this study emphasizes operational realism in both domains, offering a basis for future work that could bridge the two through power-aware or hybrid simulation frameworks.

3. METHODOLOGY

This project investigates how AI can improve both the performance and energy efficiency of 5G networks through two complementary tasks: MCS prediction and energy consumption modeling. By using simulated data to explore real-time link adaptation and real-world data to model energy usage patterns, the work connects protocol-level decisions with their sustainability implications. Although each model is developed independently, together they offer a systems-level perspective on how intelligent, adaptive mechanisms can support smarter, greener 5G deployments.

3.1 Link Adaptation / MCS Prediction

The link adaptation part of the study was conducted using the 5G-LENA simulator, which runs on top of the ns-3 framework. For simplicity and to maintain a controlled environment, the setup includes a single gNB serving four UEs. The gNB is placed at the center of a standard hexagonal cell, which is commonly used in 5G simulations to reflect realistic cell coverage. The scenario was designed for downlink traffic only, aligning with the standard practice in MCS selection studies, where decisions are made at the gNB.

The simulation uses a 3GPP-compliant channel model under an Urban Macro (UMa) scenario. The 3GPP channel model was chosen because it is the recommended standard for performance evaluation by the 3rd Generation Partnership Project (3GPP). The UMa scenario was selected to represent dense urban deployments, where buildings and mobility reflect real-world deployment in metropolitan areas.

The UEs are placed randomly within the coverage area and move at varying speeds to simulate realistic user behavior. This setup helps evaluate how the AI models respond to dynamic channel conditions when predicting optimal MCS values.

For traffic generation, the simulation uses a gaming traffic profile over the UDP protocol. This choice was made to reflect latency-sensitive communication, which places stricter demands on link adaptation compared to more tolerant traffic types like file downloads or video streaming.

The entire simulation was automated through a scripted setup to ensure consistency and reproducibility. To introduce variability and avoid bias, each simulation was run with 10 different random seeds, and every seed was executed across 3 separate runs. This provided a diverse dataset, capturing a range of mobility patterns and channel fluctuations, which was essential for training the MCS prediction models.

Parameter	**Description**
Simulator	5G-LENA (NR-v4.0)
ns-3 Version	ns-3.44
Channel Model	3GPP Urban Macro (TR 38.901), NLOS condition
Frequency	30.5 GHz
Bandwidth	100MHz
Number of gNBs	1 (center of hexagonal layout)
Number of UEs	4
Mobility Model	Random Walk with variable speed (UEs repositioned randomly for each run)
Traffic Type	Gaming traffic over UDP
Scheduling Algorithm	Default PDSCH RR Scheduler
CQI Reporting	Event-driven
MAC Stats Logging Interval	0.01 seconds
Output Files	NrDlMacStats.txt,DlDataSinr.txt, RxedGnbMacCtrlMsgsTrace.txt
Simulation Duration	200 seconds (real time per run)
Simulation Runs	10 different seeds × 3 runs = 30 runs total
Automation	bash script
OS & Toolchain	Ubuntu 20.04 LTS, g++-11.4.0

Table 1. Simulation Configuration

Running Simulation: To enable AI-driven link adaptation, a custom 5G-LENA simulation scenario was designed with parameters outlined in Table 1. The setup includes configurable aspects such as carrier frequency, channel model, and user mobility. UE positions were randomized for each run, with mobility modeled using random speeds to capture diverse propagation conditions. Each simulation was run with 10 different seeds, repeated thrice for variability, resulting in over 20,000 data points.

To keep the dataset focused yet informative, only key metrics relevant to MCS prediction were logged. Specifically, the system recorded MAC layer statistics, SINR values, and CQI reports. MAC statistics were sampled every 0.01 seconds to capture fine-grained changes in MCS and throughput. SINR and CQI values, in contrast, were logged in response to packet transmission events governed by the scheduler. This selective logging ensures the dataset remains compact while preserving the critical indicators needed to model the relationship between channel conditions and modulation strategy.

Post-Processing: All preprocessing was performed using Python, leveraging Pandas and NumPy for structured time-series handling. The core challenge was aligning features logged at different intervals: MAC stats were recorded at fixed 0.01-second intervals, while SINR and CQI values were event-triggered. We used the MCS timestamps as anchors and matched the nearest CQI and SINR values within a defined window.

Retransmissions were considered explicitly. For each MCS assignment, we counted the number of times it was repeated and treated this as a feature—capturing how often the scheduler had to resend data under poor conditions. This provided a useful proxy for link reliability.

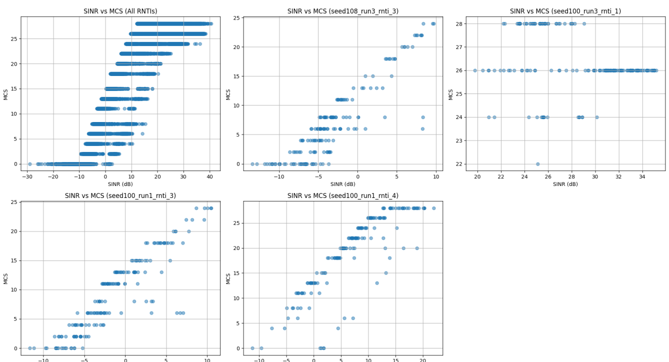


Fig. 1. SINR Vs MCS

Each run, seed, and RNTI combination was treated as a unique instance. This preserved diversity across mobility paths, channel states, and user-specific variations, rather than grouping sequences into broader batches.

Visualizations were generated to assess signal dynamics and validate pre-processing. Heatmaps of the MCS distribution revealed an uneven class presence, scatter plots between SINR and MCS showed partial correlation, and burst patterns in CQI confirmed HARQ-triggered feedback variability.

After cleaning and filtering out incomplete or noisy sequences, the final dataset consisted of more than 20,000 data points, ready for model training and evaluation.

AI Modeling: To capture the dynamics of link adaptation, a series of supervised learning models were developed, beginning with a CNN-based classifier trained on time-series sequences of SINR. Each sample consisted of 10 consecutive SINR values used to predict the MCS in the next time slot. The dataset, consisting of over 20,000 labeled instances, was split into 70% training and 30

The base CNN architecture was influenced by Zhang et al. (2024), who proposed framing MCS prediction as a classification problem grounded in recent SINR context. Following this direct prediction approach, the model operates on sliding windows of SINR values without relying on explicit channel feedback, simulating a real-time adaptation mechanism. All models were implemented in PyTorch and trained using the Adagrad optimizer (learning rate 0.01, batch size 64) over 50 epochs, with cross-entropy loss as the objective function.

To mitigate the significant imbalance across MCS classes, a weighted random sampler was employed during training, assigning higher importance to underrepresented categories.A basic hyperparameter sweep was performed over 12 sampled configurations, revealing that a CNN model with 128 convolutional filters, 8-dimensional RNTI embeddings, no dropout, and a learning rate of 0.005 achieved the highest validation accuracy around 55.3%.

In addition to the base CNN, further variants were explored including models that incorporated CQI values, encoded RNTI identifiers, and masking strategies to reflect incomplete channel feedback. Finally, a recurrent architecture using LSTM layers was tested to capture longer-term temporal dependencies in SINR behavior. This allowed comparison between convolutional and sequential memory-based approaches to time-series classification under volatile 5G radio conditions.

Evaluation & Validation: The trained models were evaluated using standard classification metrics, focusing primarily on overall accuracy and class-wise performance. A 30% validation split was held out during training to assess generalization. For interpretability, a confusion matrix was generated to visualize misclassification across MCS levels, along with per-class precision, recall, and F1 score graphs. These tools helped identify which MCS categories were more difficult to predict, especially under-represented or infrequent values. All results reported in the next section reflect validation performance after 50 training epochs.

3.2 Part II: Energy Prediction Model

Dataset Selection and Preprocessing: To investigate energy usage patterns in 5G systems, this study uses the 5G-Energy Consumption Dataset Nadia Triki (2025), sourced from Kaggle. The dataset contains 92,629 entries across 1019 unique base stations (BS), labeled from B_0 to B_1018. Each record captures hourly energy consumption for a base station over time, spanning the following attributes:

- Timestamp (Time) – in YYYYMMDD HHMMSS format
- BS ID (BS) – unique identifier for each base station
- Energy – measured energy consumption in watt-hour

- Load – traffic load on the base station
- ESMODE – operational energy-saving mode indicator
- TXpower – transmission power level

This large-scale, real-world dataset offers high spatial diversity and consistent temporal granularity. Although it lacks configuration details like frequency or bandwidth, its size and operational fidelity make it suitable for modeling coarse-grained energy trends, especially as they relate to network load and transmission behavior. It complements the simulation-based analysis by grounding the study in measurable power consumption patterns.

Preprocessing: Preprocessing was performed using Python libraries such as pandas and NumPy. The Time field was converted to datetime format to extract features like hour, weekday, and weekend flag. Base station IDs (BS) were one-hot encoded, resulting in 1,019 binary features.

No missing values were found in key features (Energy, load, TXpower). Initial visualizations showed strong positive correlation between energy use and both load and TXpower, confirmed by scatter plots and a correlation heatmap. Boxplots revealed energy-saving mode (ESMODE) effects, and time-series plots highlighted base station-level variability. A heatmap of high-variance BSs showed notable temporal diversity.

These patterns validated the relevance of both temporal and operational features for modeling.

Modelling and Evaluation: Given the limited parameters of the publicly available dataset, this phase focused on baseline modeling and reproducing select setups from Chen and et al. (2024)'s competition work. The initial models included Linear Regression and Logistic Regression, providing a baseline for energy consumption estimation.

Inspired by Chen and et al. (2024), further experiments incorporated masking and non-masking strategies to assess the effect of data sparsity and hidden feature exposure. Although the original competition dataset contained more granular data, the available subset still enabled a comparative analysis.

The evaluation of the model was performed using the mean absolute percentage error (MAPE), and the results were visualized through bar plots. While the analysis is preliminary, it establishes a foundation for follow-up studies on energy estimation using low-resolution, operator-level data.

4. RESULTS & DISCUSSION

4.1 Link Adaptation

Model	Input Features	Accuracy
CNN (baseline)	SINR	52.3%
CNN + CQI	SINR + CQI	53.71%
CNN + Masking	SINR + CQI + Masking	54.11%
LSTM (best)	SINR + CQI	55.15%

Table 2. Link Adaptation Model Comparison

Across all trained models, classification accuracy ranged from approximately 52% to just above 55%, with the LSTM-based model achieving the best performance. As shown in Table2, even modest architectural variations—such as incorporating CQI or applying input masking—produced incremental gains, though none significantly outperformed the LSTM baseline. These results reflect the challenge of modeling MCS prediction under dynamic channel conditions using using data with significant class imbalance and capturing a specific range of dynamic channel conditions.

Analysis of the confusion matrices (A.2 in the appendix) and per-MCS performance plots reveals a pronounced skew in predictive accuracy. MCS levels such as 0, 24, 26, and 28 were classified more reliably than others. These values are also the most frequent in the dataset, leading to stronger learned associations. Despite using weighted sampling during training to counteract class imbalance, rarer MCS levels remained poorly predicted—and in some cases, were never predicted at all—due to their near absence in the training data.

This skew is not merely a byproduct of simulation artifacts; in practice, 5G systems often prefer higher MCS levels to maximize throughput under favorable conditions. Thus, real-world datasets are also likely to reflect similar distributions. To improve class coverage and generalization, future work could incorporate a wider range of simulated scenarios, including more diverse user positions, traffic types, or denser deployments. These enhancements would support better exposure to underrepresented MCS classes, but were outside the computational scope of this study.

Still, the comparative model performance affirms that sequence-based models are more adept at learning from temporal channel variations. This aligns with the findings of Stenhammar et al., who showed that recurrent networks like LSTM and GRU outperform CNNs, MLPs, and even Transformers for fast-fading wireless channels. Their work reinforces the conclusion that link adaptation, being inherently time-sensitive, benefits from models that can track sequential dynamics across time slots.

4.2 Energy Prediction

Initial attempts at energy consumption prediction using basic linear regression resulted in high error values, with a MAPE exceeding 30% and poor generalization. The model struggled to capture non-linear relationships in the data, despite a moderately low MAE of 7.5.

To improve performance, a simple three-layer neural network was implemented, achieving a MAPE of 16.9% after 100 training epochs. This marked a substantial improvement, though it still fell short of the desired predictive precision.

Following the approach proposed by Chen and et al. (2024), a masked variant of the model was evaluated. This version incorporated structured learning over masked input segments and achieved a MAPE of 5.5% within just 31 epochs, demonstrating significantly faster convergence and higher accuracy. These results align closely with Chen and et al. (2024)'s reported 4.98%, though their models benefited from a more feature-rich dataset—including detailed base station configurations and environment-aware signals—which were not available in the public version used here.

The findings suggest that, even with minimal feature engineering, energy consumption in 5G base stations can be accurately estimated using learned patterns in load and transmission power. If enhanced with richer telemetry, such models could support intelligent decisions that jointly optimize throughput and energy efficiency. For example, since higher MCS levels generally correlate with increased transmission power, an ideal adaptation strategy would balance spectral efficiency with consumption, especially in dense or power-constrained deployments.

5. CONCLUSION

This study explored AI-driven strategies for two core challenges in 5G networks: link adaptation through MCS prediction, and energy consumption modeling at the base station level. Using custom 5G-LENA simulations and a public energy dataset, the analysis applied a range of supervised learning techniques—from CNNs and LSTMs to regression and masked neural networks.

For link adaptation, results showed that sequence-aware models, particularly LSTMs, outperformed CNNs and static classifiers, albeit by a modest margin. Prediction accuracy peaked at just over 55%, constrained largely by dataset imbalance and the naturally skewed MCS usage patterns in realistic transmissions. High-frequency MCS values like 0, 24, and 28 were predicted reliably, while rare values remained largely undetected. Despite using a weighted sampler, these challenges persisted—highlighting the need for larger, more diverse simulation datasets or synthetic augmentation for robust model training.

In parallel, the energy modeling task demonstrated that even simple neural networks could meaningfully improve predictive performance over linear baselines. When inspired by the masked modeling strategy proposed by Chen and et al. (2024), the model achieved a MAPE of 5.5%, closely aligning with reported benchmarks. This suggests that energy-aware decision-making in RAN could be made feasible with relatively lightweight models, particularly when enriched by more granular telemetry.

Together, the two parts of this work offer a preliminary framework for dual-objective optimization in 5G: selecting MCS values that not only adapt to changing channel conditions but also consider their impact on power consumption. While this research does not yet unify these two modeling paths, it lays the foundation for future research that does—potentially using reinforcement learning to make dynamic, energy-efficient link adaptation decisions in real-time.

5.1 Future Work

In the immediate future, several improvements are possible within the scope of this project. More refined feature engineering, such as incorporating time-based patterns or CQI deltas, might lead to more accurate predictions. Models like GRU, which have shown strong performance in related time-series tasks, could be implemented as a follow-up to the current LSTM setup. A better dataset—ideally with more varied traffic patterns, denser deployment, or targeted simulation of underrepresented MCS classes—would help mitigate class imbalance and improve overall generalization. These are all tasks that can realistically be attempted within the next couple of months and would strengthen the practical outcomes of this work.

In the longer term, there is room to explore joint optimization approaches that consider both throughput and energy efficiency, something that becomes more relevant when transitioning from simulation to deployment. Models such as MiniROCKET or HDC-based classifiers, especially those designed for low-complexity, time-sensitive tasks, could be tested as lightweight alternatives to deep learning. Better energy models could also be developed using additional features, especially if richer datasets become available. Finally, more systematic hyperparameter tuning and repeated validation runs would help in building more stable and reproducible results—something worth aiming for in future iterations.

Finally, this work underscores a broader challenge in AI-driven 5G research: the lack of accessible, realistic datasets that reflect the variability and complexity of live network environments. While simulation provides a useful starting point, it cannot fully capture the noise, delays, and unexpected interactions present in actual deployments. For link adaptation and energy modeling to evolve into deployable solutions, the field needs more open datasets, richer in both features and context. That, more than anything, may determine how far these AI techniques can really go.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my supervisor, Professor Evgeny Osipov, for his guidance, expertise, and patience throughout this project. I am deeply thankful to the Erasmus Mundus Green Networking and Cloud Computing program and the GDL coordinators for making this opportunity possible. Special thanks to Effnet for their support and technical insights. I am especially grateful to my family and friends for their unwavering encouragement

REFERENCES

Abubakar, A.I., Omeke, K.G., Ozturk, M., Hussain, S., and Imran, M.A. (2020). The role of artificial intelligence driven 5g networks in covid-19 outbreak: Opportunities, challenges, and future outlook. *Frontiers in Communications and Networks*, 1. doi: 10.3389/frcmn.2020.575065.

Balasubramanian, B. et al. (2021). Ric: A ran intelligent controller platform for ai-enabled cellular networks. *IEEE Internet Computing*, 25(2), 7–17. doi: 10.1109/MIC.2021.3062487.

Chen, T. and et al. (2024). Modelling the 5G Energy Consumption using Real-world Data: Energy Fingerprint is All You Need. doi:10.48550/arXiv.2406.16929. Accessed: 2025-05-27.

Elgabroun, H.M.A.I. (2019). *Machine Learning Technique for Uplink Link Adaptation in 5G NR RAN at Millimeter Wave Frequencies.* Master's thesis, Lund University. Http://lup.lub.lu.se/student-papers/record/9118249.

Gigayasa Wireless (2025). Link Adaptation — 5G Toolkit R24a Documentation. Accessed March 4, 2025.

Herath, J.D., Seetharam, A., and Ramesh, A. (2019). A deep learning model for wireless channel quality prediction. In *ICC 2019 - IEEE International Conference on Communications*, 1–6. IEEE, Shanghai, China. doi:10.1109/ICC.2019.8761934. Https://ieeexplore.ieee.org/document/8761934/.

Kojima, S., Maruta, K., and Ahn, C.J. (2019). Adaptive modulation and coding using neural network based snr estimation. *IEEE Access*, 7, 183545–183553. doi:10.1109/ACCESS.2019.2946973. Https://ieeexplore.ieee.org/document/8865038/.

Lee, H., Jang, Y., Song, J., and Yeon, H. (2021). O-ran ai/ml workflow implementation of personalized network optimization via reinforcement learning. In *2021 IEEE Globecom Workshops (GC Wkshps)*, 1–6. doi:10.1109/GCWkshps52748.2021.9681936.

Li, Y., Xu, Y., Li, G., Wang, X., Xu, Y., Liu, S., and Yue, L. (2024). Turning adversity into advantage: Robust mcs selection utilizing the uncertainty of channel prediction neural networks. *IEEE Wireless Communications Letters*, 13(10), 2632–2636. doi:10.1109/LWC.2024.3393939. Https://ieeexplore.ieee.org/abstract/document/10654736.

Nadia Triki (2025). 5g-energy consumption dataset. Https://www.kaggle.com/datasets/nadiatriki/5g-energy-consumption-dataset (Accessed: May 20, 2025).

Norolahi, J. and Azmi, P. (2023). A machine learning based algorithm for joint improvement of power control, link adaptation, and capacity in beyond 5g communication systems. *Telecommunication Systems*, 83(4), 323–337. doi:10.1007/s11235-023-01017-1.

Oh, J.E., Jo, A.M., and Jeong, E.R. (2023). Mcs selection based on convolutional neural network in mobile communication environments. In *2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN)*, 684–686. doi:10.1109/ICUFN57995.2023.10201063. Https://ieeexplore.ieee.org/abstract/document/10201063.

Parsa, A., Moghim, N., and Salavati, P. (2022). Joint power allocation and mcs selection for energy-efficient link adaptation: A deep reinforcement learning approach. *Computer Networks*, 218, 109386. doi: 10.1016/j.comnet.2022.109386.

Rezazadeh, F. et al. (2023). On the specialization of fdrl agents for scalable and distributed 6g ran slicing orchestration. *IEEE Transactions on Vehicular Technology*, 72(3), 3473–3487. doi:10.1109/TVT.2022.3218158.

Sahu, R. and Sahu, V. (2024). An energy-efficient algorithm for resource allocation in h-cran (ee h-cram) for 5g networks. *Wireless Personal Communications*, 138(3), 1483–1499. doi:10.1007/s11277-024-11545-x.

Seeram, S.S.S.G.S. (2022). *Link adaptation in 5G Networks: Reinforcement Learning framework based approach*. Master's thesis, KTH Royal Institute of Technology.

Stenhammar, O., Fodor, G., and Fischione, C. (2024). A comparison of neural networks for wireless channel prediction. *IEEE Wireless Communications*, 31(3), 235–241. doi:10.1109/MWC.006.2300140.

Tran, N.B.L. et al. (2024). Ai-driven rapps for reducing radio access network interference in real-world 5g deployment. In *IEEE INFOCOM 2024 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 1–2. doi: 10.1109/INFOCOMWKSHPS61880.2024.10620743.

Tsipi, L., Karavolos, M., Papaioannou, G., Volakaki, M., and Vouyioukas, D. (2024). Machine learning-based methods for mcs prediction in 5g networks. *Telecommunication Systems*, 86(4), 705–728. doi:10.1007/s11235-024-01158-x. Https://link.springer.com/article/10.1007/s11235-024-01158-x.

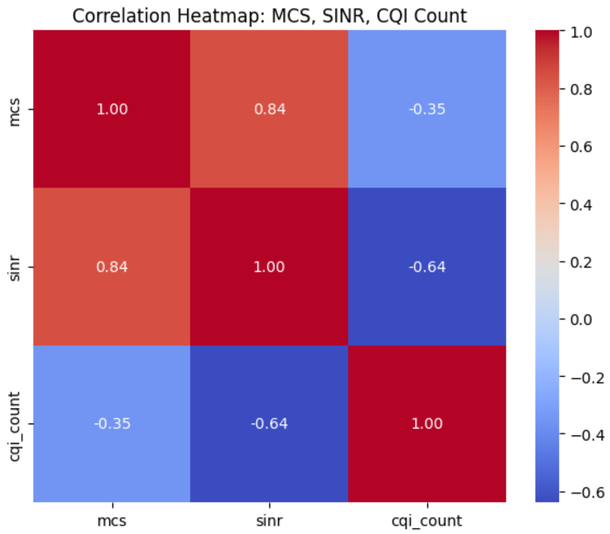
Varshney, R., Gangal, C., Sharique, M., and Ansari, M.S. (2023). Deep learning based wireless channel prediction: 5g scenario. *Procedia Computer Science*, 218, 2626–2635. doi:10.1016/j.procs.2023.01.236. Https://www.sciencedirect.com/science/article/pii/S1877050923

Yun, W., Hong, S., and Jeong, E. (2024). Gru-based mcs selection for uav communication in 5g environment. In *2024 Fifteenth International Conference on Ubiquitous and Future Networks (ICUFN)*, 472–475. doi:10.1109/ICUFN61752.2024.10625063. Https://ieeexplore.ieee.org/document/10625063.

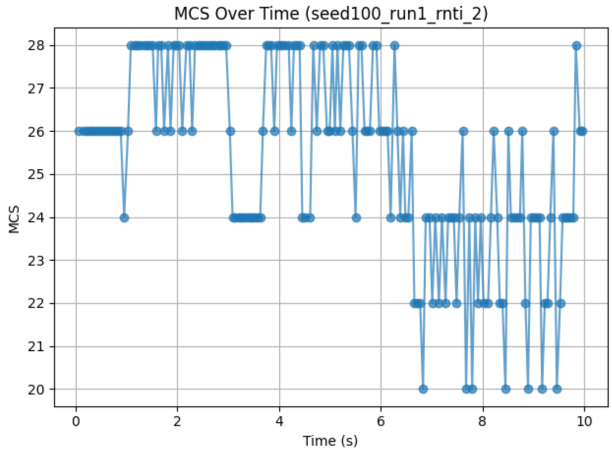
Zhang, C.K., Hossain, E., and Kim, D.I. (2018). Deep convolutional neural networks for link adaptations in mimo-ofdm wireless systems. *IEEE Transactions on Wireless Communications*, 17(4), 2621–2633. doi:10.1109/TWC.2018.2804958. Https://ieeexplore.ieee.org/document/8540019.

Zhao, Z., Feng, C., Yang, H.H., and Luo, X. (2020). Federated-learning-enabled intelligent fog radio access networks: Fundamental theory, key techniques, and future trends. *IEEE Wireless Communications*, 27(2), 22–28. doi:10.1109/MWC.001.1900370.

Appendix A. TABLES & IMAGES



(a) Sample Correlation Heat map

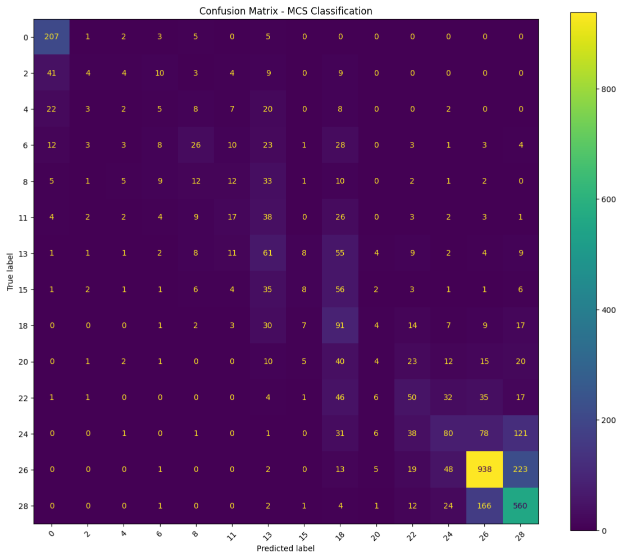


(b) Sample for MCS vs Time

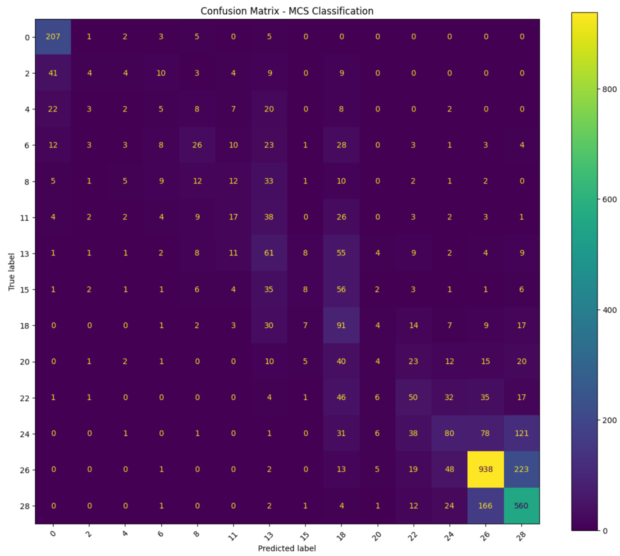
Fig. A.1. Samples for visualisation

Table A.1. Comparison of Datasets in Recent MCS/Channel Prediction Studies

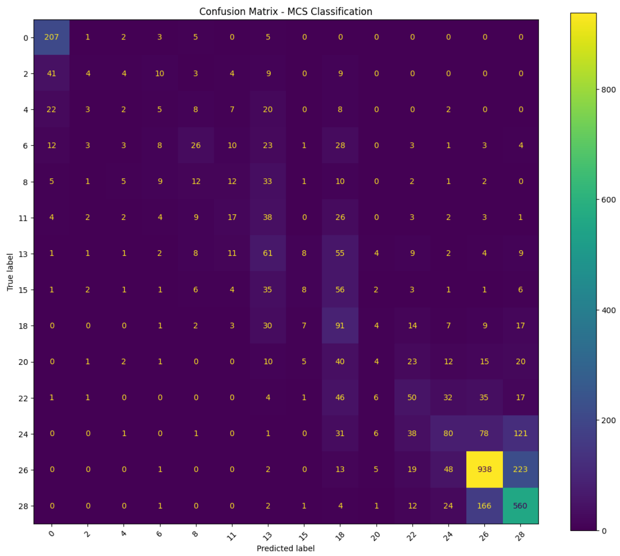
Study	**Data Method**	**Scenario**	**Realism**	**Tool / Company**
Varshney et al. (2023)	NYUSIM-based mmWave sim	28 GHz Urban Macro	High	NYUSIM / NYU
Zhang et al. (2018)	MIMO-OFDM, SNR sweep	Synthetic system-level	Medium-Low	–
Herath et al. (2019)	Real RSS from 5 networks	LTE, WiFi, Zigbee, etc.	Real-world	CRAWDAD / G5
Li et al. (2024)	RSSI via spectrum analyzer	2.4 GHz RSSI trace	Real-world	Spectrum Analyzer
Seeram (2022)	MATLAB 5G testbed	SU-MIMO, mobility, HARQ	Medium-High	Huawei
Oh et al. (2023)	Rayleigh/Rician in MATLAB	5G NR-like setup	Medium	MATLAB/TensorFlow
Elgabroun (2019)	mmWave sim (UL)	28 GHz, walking UE	High	Ericsson tool
Stenhammar et al. (2024)	3GPP TDL-A model	2 GHz NLOS, mobile UE	Medium-High	3GPP / –
Tsipi et al. (2024)	Ray-tracing + LTE overlay	Urban 2.1 GHz NSA	High	Altair FEKO
Yun et al. (2024)	5G NR sim, Rayleigh/Rician	UAV, 3.6 GHz, 100 MHz	Medium	MATLAB/TensorFlow



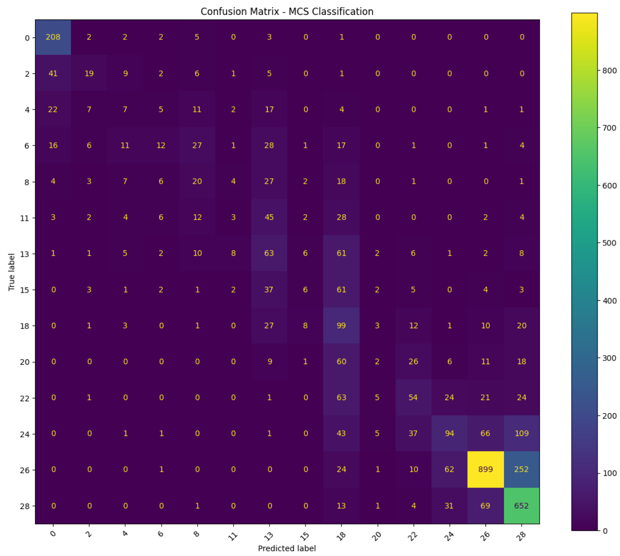
(a) SINR only



(b) SINR & CQI

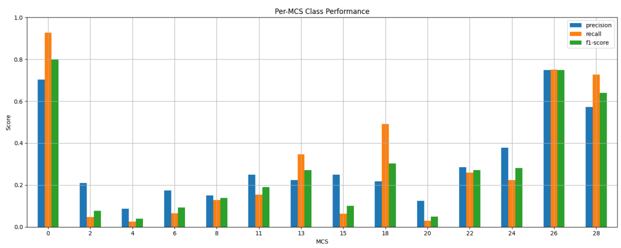


(c) Masking

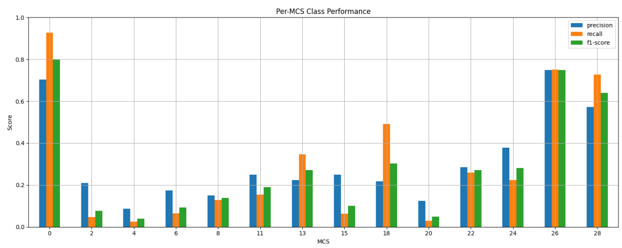


(d) LSTM

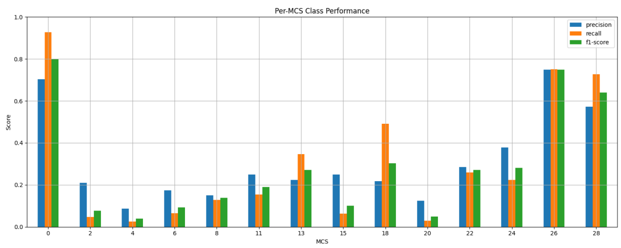
Fig. A.2. Confusion matrices for MCS.



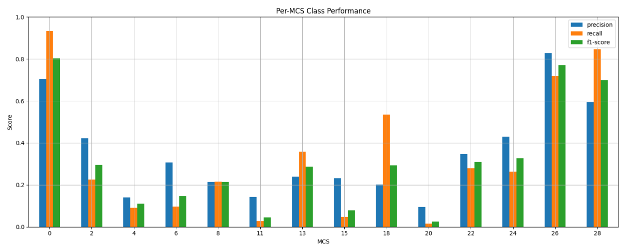
(a) SINR only



(b) SINR & CQI



(c) Masking



(d) LSTM

Fig. A.3. per MCS Class performance for MCS Prediction