From Statistics to Deep Learning: Forecasting Mobile Throughput

Estêvão J. Santos[1] and Tiago A. Almeida[1]

Universidade Federal de São Carlos (UFSCar)
estevaojs@estudante.ufscar.br, talmeida@ufscar.br

Abstract. Accurate download throughput prediction is critical for adaptive resource management and QoS in 5G networks, particularly under high user mobility. This work systematically investigates two key design choices in time series forecasting: (i) local versus global models and (ii) the inclusion or exclusion of external covariates. We evaluate statistical, machine learning, and deep learning methods on real-world 5G data, where throughput is predicted using channel quality metrics and user speed as potential covariates. Experimental results show that global, tree-based ensembles like LightGBM achieve the best trade-off between accuracy, robustness, and efficiency. Furthermore, we found that the explored network quality covariates were insufficient to consistently improve performance for this complex task. All source code is available at: `https://github.com/ejs94/5g-forecasting`.

CCS Concepts: • **Computing methodologies** → **Time series analysis**; *Neural networks*; • **Networks** → **Mobile networks**.

Keywords: time series forecasting, 5G networks, throughput prediction, machine learning, mobility, quality of service

1. INTRODUCTION

With the deployment of 5G networks and the gradual phase-out of 4G systems, mobile broadband usage has surged due to a growing user base and expanded coverage. While 4G operates in the 700 MHz to 2.6 GHz range, 5G employs both sub-6 GHz and millimeter-wave (mmWave) bands, substantially increasing spectral efficiency and enabling data rates well beyond those of previous generations [Ghosh et al. 2019].

Despite these advances, 5G deployment remains uneven, especially in suburban and rural areas. Moreover, higher-frequency bands suffer from greater signal attenuation and sensitivity to obstructions, which significantly limits coverage compared to legacy technologies [Boutiba et al. 2021]. In this scenario, ensuring efficient and adaptive network behavior becomes increasingly difficult, particularly under varying conditions of traffic, topology, and mobility.

Accurate throughput prediction emerges as a crucial component for adaptive resource management and Quality of Service (QoS) provisioning in such networks. Forecasts of near-future traffic loads can drive optimizations like opportunistic scheduling, energy-aware control, and dynamic handover management. These capabilities are vital for supporting latency-sensitive services such as real-time video streaming and autonomous vehicles [Santos et al. 2020].

Throughput is inherently influenced by radio channel conditions, typically quantified by metrics such as Reference Signal Received Power (RSRP), Reference Signal Received Quality (RSRQ), Signal-to-Noise Ratio (SNR), and Channel Quality Indicator (CQI) [Raca et al. 2020]. While RSRP measures signal strength, RSRQ and SNR describe signal quality and interference, directly affecting the modulation and coding schemes the system can use. CQI, in turn, is the user device's recommendation on the data rate the link can sustain, reflecting both signal conditions and radio link capacity.

However, this relationship is volatile. User mobility introduces a dynamic temporal dimension that significantly affects throughput estimation. In high-speed scenarios, frequent handovers, body shadowing, and abrupt SNR fluctuations impair stability [Narayanan et al. 2020]. Even at walking speeds, subtle variations in orientation or movement direction may obstruct the line-of-sight path—especially in mmWave bands—degrading the quality of service. These mobility-induced fluctuations render static performance models insufficient and underscore the need for time-aware, data-driven predictive approaches.

Time series forecasting—the task of predicting future values from past observations—is particularly suitable for addressing this challenge. Unlike traditional supervised learning tasks, forecasting requires preserving temporal dependencies, explicitly modeling uncertainty, and handling phenomena like non-stationarity, seasonality, and structural breaks [Hyndman and Athanasopoulos 2021]. To ensure reliable evaluation in such settings, walk-forward validation is commonly adopted [Hewamalage et al. 2023], as it respects the temporal order and mitigates the risk of data leakage from future observations.

The evolution of forecasting methods has been strongly influenced by large-scale empirical benchmarks, such as the M-series competitions [Hyndman 2020], which have demonstrated that forecasting accuracy depends not only on the choice of modeling paradigm—statistical or machine learning—but also on the model's functional capabilities, particularly its ability to incorporate external covariates [Makridakis et al. 2022]. In the context of mobile networks, such covariates may include user velocity, signal quality metrics, time of day, and cell-specific identifiers. Moreover, as highlighted by [Januschowski et al. 2020], the distinction between statistical and machine learning approaches has become increasingly ambiguous, as both fields converge through shared techniques and the emergence of hybrid or unified modeling frameworks.

In this context, a growing body of research has examined two orthogonal design choices in forecasting architectures: (i) local models, which are trained independently for each time series, versus global models, which learn shared patterns across multiple series [Yingjie and Abolghasemi 2024]; and (ii) models that do or do not incorporate external covariates. In this work, we explicitly investigate the second dimension—whether or not to include covariates—by evaluating models in both configurations.

Selecting the appropriate forecasting strategy in this setting thus requires careful consideration of network dynamics, feature relevance, model scalability, and application-specific constraints. This work investigates and compares such methods for short-term throughput prediction, aiming to support traffic-aware optimization mechanisms in next-generation mobile networks.

1.1 Objectives and Research Questions

This study addresses key gaps in applying time series forecasting to mobile networks, especially in high-mobility scenarios. The main objectives are to (i) present and evaluate time series forecasting methods for download throughput prediction in 5G networks; (ii) compare local and global models to assess generalization, scalability, and sensitivity to mobility and signal variations; and (iii) support proactive radio resource management by leveraging throughput predictions to enhance Quality of Service (QoS).

The research questions guiding this investigation are: (RQ1) Which forecasting methods are most effective under varying network and mobility conditions? (RQ2) What are the comparative advantages and limitations of local versus global modeling approaches?

1.2 Contributions

This work contributes to predictive resource management in cellular networks by offering comprehensive benchmarking of statistical, local, and global learning models for throughput forecasting; actionable insights for model selection based on mobility, signal quality, and interpretability; and empirical evidence supporting the use of time series predictions for proactive QoS mechanisms in 5G.

2. RELATED WORKS

The study by [Batool et al. 2024] explores throughput prediction in 5G using deep learning models, comparing LSTM, BiLSTM, ANN, and Random Forest. Although BiLSTM achieves the lowest MSE (0.00009), ANN obtains the highest R^2 score (0.36), revealing a trade-off between prediction accuracy and explained variance. The authors justify the use of LSTM-based models based on their strong capacity to model temporal dependencies in network traffic. By incorporating both radio metrics (e.g., SNR, RSRQ, CQI) and contextual data (e.g., location, speed), the study reinforces the relevance of multi-source features for modeling throughput. These insights are particularly valuable for applications involving both local and global models, supporting adaptive and cost-efficient network management under realistic 5G scenarios.

In their study on 4G LTE networks, [Elsherbiny et al. 2020] conducted a comparative analysis of various methods for throughput prediction to support proactive resource allocation. Their findings highlight the superiority of classical Machine Learning ensembles, with a Random Forest model achieving the highest accuracy (R^2 of approx. 0.78), outperforming not only SVR and KNN but also a deep learning-based LSTM. Notably, the traditional statistical model ARIMA also demonstrated competitive performance, reinforcing that model complexity does not always guarantee better results on this type of dataset. The study concluded that beyond primary signal metrics (RSRP, RSRQ, etc.), contextual features such as GPS location, speed, and time of day were critical predictors, underscoring the effectiveness of ensemble methods that can leverage such diverse inputs.

The Lumos5G framework by [Narayanan et al. 2020] tackles throughput prediction challenges in 5G mmWave networks, characterized by highly variable signal quality. The study shows that traditional 4G-based, location-only approaches fail to accurately predict 5G throughput due to environmental sensitivity and mobility factors. Lumos5G leverages a contextual, machine learning–based approach using Gradient Boosted Decision Trees (GDBT) and Sequence-to-Sequence (Seq2Seq) models. Features are grouped into location, mobility, tower characteristics, and connection state. This multi-faceted contextual information significantly improves throughput prediction accuracy, demonstrating that location alone is insufficient for dynamic urban 5G environments.

The study by [Sharma et al. 2025] addresses 5G throughput prediction as a classification problem, proposing the Enhanced Sequential Decision Tree (ESDT) model. This approach achieved superior results across multiple metrics, including 92.5% accuracy, 93.5% recall, 92.7% F1-score, and 97% AUC. A key contribution is the hybrid feature selection strategy, combining embedded methods with forward selection, which, along with effective preprocessing, improved model performance. The most relevant features included geolocation, movement direction, connection status, speed, compass heading, and LTE metrics such as RSRP and RSSI. The results highlight the importance of careful data preparation and demonstrate strong generalization across multiple 5G datasets.

Finally, the work of [Yeaser and Hassan 2025] investigates throughput prediction in 5G NR V2X communications, emphasizing its importance for supporting critical vehicular applications. The authors propose a hybrid deep learning model combining LSTM and GRU architectures, trained on real-world data. Their analysis revealed that the model achieved higher accuracy in predicting downlink throughput, while uplink prediction was less accurate due to the sporadic nature of uplink transmissions, which occur only when needed. Input features included SNR, speed, RSRP, and RSRQ, capturing key effects from mobility, path loss, and interference. These findings highlight the potential of hybrid models to provide reliable throughput prediction in dynamic vehicular environments.

3. METHODOLOGY

This section details the methodology for predicting download throughput using channel quality metrics as covariates. The following sections cover the dataset used, the preprocessing steps applied to the data, the performance measures employed to evaluate the models, and the experimental protocol detailing the training and testing procedure. Finally, the results achieved by the prediction models are also discussed, providing insights into their performance and implications.

3.1 Dataset

This study uses a dataset introduced by [Raca et al. 2020], which includes download throughput measurements collected under diverse mobility and application scenarios. The dataset comprises 83 traces, totaling over 3,000 minutes of data, with one-second resolution for data collection. The traces reflect real-world usage patterns and include both static and driving conditions, as well as application types such as video streaming and file downloads.

For the throughput prediction task, we considered download throughput as the target variable. As covariates, we selected RSRP, RSSI, RSRQ, SNR, CQI, and UE speed, given their strong relevance in characterizing radio conditions and mobility. This choice is supported by recent studies by [Sharma et al. 2025] and [Narayanan et al. 2020], which demonstrate that combining signal quality metrics with mobility features improves throughput prediction across various modeling approaches. These variables are also practical, as they are commonly available at the UE level in real deployments.

All 83 time series were analyzed for trend and seasonality. However, due to the fine temporal granularity (one-second intervals), most series did not exhibit explicit trends or seasonal patterns. As a result, classical decomposition methods were not applied, and stationarity was instead assessed using the Augmented Dickey-Fuller (ADF) test, which indicated that many series were either stationary or required minimal differencing.

3.2 Preprocessing

The dataset underwent a standardized preprocessing pipeline to ensure consistency for time series modeling. Timestamps were normalized, duplicates removed, and key signal metrics (RSRP, RSSI, RSRQ, SNR, CQI) along with the throughput target were scaled via min-max normalization, following common practice [Batool et al. 2024]. Missing values were addressed through a combination of linear interpolation for internal gaps and forward/backward filling for edge cases, with zero imputation applied only when sequences were entirely missing. Although backward filling risks introducing future information, this was restricted to preprocessing, while walk-forward validation preserved causality during evaluation [Sharma et al. 2025].

Unlike some works that include contextual features like GPS or operator IDs [Narayanan et al. 2020], this study excluded such variables to enhance generalizability across regions. Prior studies that incorporate these features emphasize location-specific patterns [Batool et al. 2024], but here the focus remained on signal quality metrics alone, producing clean, temporally coherent inputs suitable for both local and global forecasting models.

3.3 Evaluation Metrics

To assess model performance, we utilized the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE). These metrics are widely adopted for time series forecasting as they provide a direct and interpretable measure of prediction error.

The final performance of each model was determined by averaging these metrics across all individual time series in the dataset. This average score served as the primary benchmark to compare the effectiveness of the different approaches tested: local versus global models, and models trained with versus without covariates. The necessary computations were performed using a specialized software library designed for time series analysis.

3.4 Methods

The models evaluated span classical, machine learning, and deep learning paradigms to enable a broad comparison of forecasting strategies. Tree-based models such as LightGBM and Random Forest were selected due to their robustness, ability to capture non-linear relationships, and consistent performance in prior works [Sharma et al. 2025]. LightGBM was also a key component in the winning solutions of the M5 forecasting competition [Makridakis et al. 2022]. Linear regression was included as a lightweight and interpretable baseline, commonly used in mobility-aware throughput studies [Sharma et al. 2025]. Classical statistical models—AutoARIMA and Exponential Smoothing—were selected for their competitive performance in short-term forecasting [Makridakis et al. 2022], despite their simplicity.

Deep learning models (Block RNN, N-BEATS, Transformer) were included for their ability to model temporal dependencies and handle complex patterns. While N-BEATS was noted in the M5 competition for its potential [Makridakis et al. 2022], the others were chosen based on their success in broader time series tasks. We also evaluated global models with covariates to explore the impact of signal-related features. Although accuracy gains were modest in our case, previous studies highlight their importance for improving forecast precision [Narayanan et al. 2020]. Overall, this selection balances interpretability, performance, and scalability for the 5G throughput prediction task.

3.5 Experimental Protocol

To evaluate model generalization under realistic conditions, we adopted the walk-forward validation strategy, a method that respects the temporal dependency of time series data. Each model was initially trained on the first 80% of the data, with the remaining 20% reserved for the validation phase. During this phase, forecasts were generated iteratively over a fixed 10-step prediction horizon, with the validation window advancing sequentially to assess performance across all possible points in the reserved data.

A critical distinction was made in the validation protocol for local and global models. Local models, typically statistical, underwent an adaptive retraining strategy, where they were recalibrated at each step. While this allows for continuous adjustment, it makes them inherently susceptible to concept drift, as their performance relies on stable statistical patterns that can shift abruptly. In contrast, global models were assessed using a fixed model strategy, trained only once. Their theoretical advantage lies in learning generalized, latent patterns from a diverse corpus of series, providing greater intrinsic robustness against localized concept drift.

We selected this comprehensive approach for its robustness in evaluating forecasting models across different operating regimes, such as changes in mobility or signal quality. By assessing performance over multiple, chronologically ordered windows, walk-forward validation offers a realistic and rigorous measure of a model's predictive power in a real-world context.

The experiments were conducted on a standalone server providing sufficient computational power, equipped with multi-core Intel Xeon CPUs and an NVIDIA Tesla P100 GPU. To optimize performance, resource allocation was tailored to the model architecture: the GPU was dedicated to accelerating the training of deep learning models, while statistical and ensemble methods were executed on the CPU.

Finally, due to the large number of experiments, we kept the parameters of all methods with their default values from the Darts library[1].

4. RESULTS

The performance of the forecasting models is detailed in Figure 1 and Table I. A comprehensive statistical analysis of the error distribution (medians, quartiles, and extrema) was conducted to provide a more robust assessment than simple averages, which can be misleading as they fail to capture performance variability and outlier sensitivity.

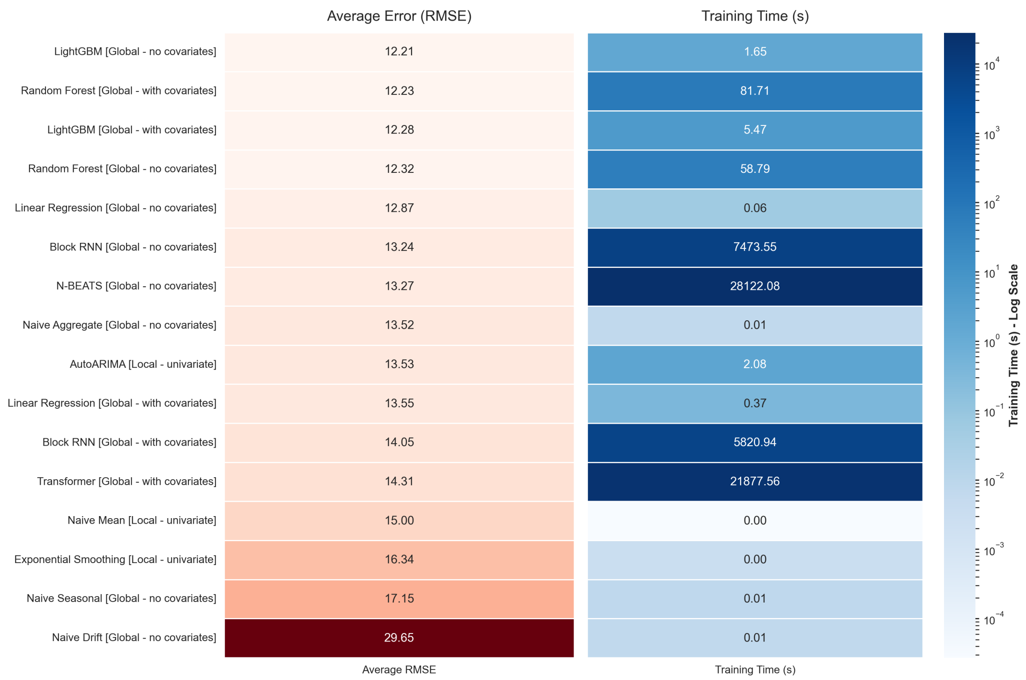

Fig. 1. Performance analysis of forecasting models, comparing Average RMSE against Training Time. This visualization highlights the clear trade-off between predictive accuracy and computational cost, where models like LightGBM demonstrate a strong balance compared to the more computationally intensive deep learning approaches. (Source: own authorship)

The analysis reveals a clear trade-off between typical accuracy and robustness. AutoARIMA achieved the lowest median RMSE (3.49) and MAE (1.40), establishing it as the most accurate model for the typical time series. However, it also exhibited poor robustness, recording the highest maximum RMSE (108.88) and MAE (94.33), indicating high sensitivity to atypical series. In contrast, LightGBM delivered strong, robust performance. While its median RMSE was slightly higher (3.59 without covariates), its maximum errors were substantially lower (RMSE < 94), demonstrating greater resilience to outliers. Furthermore, LightGBM was exceptionally efficient, with a training time of just 1.65 seconds, making it ideal for large-scale applications.

Other models showed limitations. Simple baselines like Naive Drift performed poorly and were unsuitable for this task. The inclusion of covariates yielded mixed and marginal effects, failing to provide a consistent improvement and occasionally degrading performance. Overall, LightGBM without

[1]Darts library, available at `https://unit8co.github.io/darts/`

covariates stands out as the most effective model, offering a superior balance of high accuracy, robustness, and outstanding computational efficiency, making it the most compelling choice for practical deployment.

Table I. Summary statistics of RMSE and MAE for forecasting models.

	RMSE					MAE				
Model	Min	Q1	Median	Q3	Max	Min	Q1	Median	Q3	Max
Global Models (with covariates)										
LightGBM	0.524	1.655	3.617	5.250	90.386	0.433	0.955	1.464	3.094	70.925
Transformer	0.527	1.671	3.660	5.921	108.475	0.444	0.928	1.499	3.880	80.811
Random Forest	0.650	1.670	3.678	5.158	92.014	0.496	0.908	1.521	3.103	70.471
Linear Regression	0.552	1.865	3.799	5.749	99.415	0.465	1.208	1.782	4.259	81.472
Block RNN	0.418	1.848	3.854	6.142	101.027	0.389	0.929	1.439	3.628	75.971
Global Models (without covariates)										
LightGBM	0.527	1.664	3.589	5.204	93.353	0.426	0.909	1.472	3.045	70.227
Block RNN	0.612	1.733	3.618	5.812	106.978	0.471	0.884	1.459	3.343	78.408
Random Forest	0.610	1.870	3.694	5.436	93.024	0.517	0.892	1.530	3.127	70.989
Naive Aggregate	0.001	1.756	3.796	5.509	106.373	0.001	0.712	1.408	2.843	82.978
Linear Regression	0.414	1.839	3.835	5.643	97.778	0.339	1.036	1.708	3.497	73.857
N-BEATS	0.694	1.817	3.870	5.800	102.630	0.603	1.312	1.977	3.975	79.067
Naive Seasonal	0.001	2.191	4.697	7.054	128.076	0.000	0.796	1.505	3.295	106.022
Naive Drift	0.002	3.452	8.193	12.441	217.441	0.001	1.499	2.779	5.923	184.954
Local Models (univariate)										
AutoARIMA	0.311	1.580	3.489	5.483	108.879	0.151	0.797	1.400	2.780	94.327
Naive Mean	0.665	1.662	3.777	6.326	96.785	0.550	0.835	1.491	4.708	83.510
Exponential Smoothing	0.001	1.891	4.011	6.062	122.763	0.001	0.739	1.413	2.949	98.667

5. CONCLUSION

This study compared forecasting models for 5G download throughput, evaluating the trade-offs between accuracy, robustness, and computational cost. Our analysis identifies tree-based machine learning (ML) models as the most practical solution for deployment, with the main conclusions summarized below.

RQ1: Which forecasting methods are most effective under varying network conditions?

The most effective forecasting method was identified as LightGBM, particularly when used without covariates. Its effectiveness stems not from excelling on a single metric, but from providing an optimal balance of high predictive accuracy, strong robustness against outliers, and exceptional computational efficiency. While some statistical models achieved slightly higher median accuracy, LightGBM's resilience to extreme errors and its minimal training time make it the superior and more reliable choice for dynamic and unpredictable network environments.

RQ2: What are the relative strengths and weaknesses of statistical versus ML-based approaches?

The analysis revealed distinct trade-offs between the two paradigms. The primary strength of the statistical approach, exemplified by AutoARIMA, is its ability to achieve very high accuracy on typical, well-behaved time series by fitting a specialized model. However, its main weakness is a significant lack of robustness, making it highly sensitive to noise and prone to large errors when faced with atypical data patterns.

Conversely, the strength of ML-based models like LightGBM lies in their superior robustness and ability to generalize across a diverse set of time series, reliably avoiding catastrophic failures. Furthermore, tree-based ML models are extremely computationally efficient. Their potential weakness is that their accuracy on any single, predictable series might be marginally lower than that of a perfectly tuned statistical model. Other ML models, particularly deep learning approaches, were found to be computationally intensive. In our experiment, this high cost did not translate into superior predictive performance, a result we attribute to the limited size of our training dataset, which can be insufficient for these data-hungry architectures.

Finally, a key secondary finding was that relying solely on network quality metrics as covariates was insufficient to consistently improve model performance. This aligns with the broader literature, such as [Narayanan et al. 2020], which suggests that predicting complex 5G throughput requires a richer combination of features, including context-aware parameters like location and mobility. Future work should therefore focus on integrating this more diverse set of covariates and exploring multivariate models to better capture inter-series dependencies.

Acknowledgements

We gratefully acknowledge the support provided by the Brazilian agencies FAPESP (SMARTNESS, grant 2021/00199-8) and CNPq (grant 311867/2023-5).

REFERENCES

Batool, I., Fouda, M. M., and Fadlullah, Z. M. Deep Learning-Based Throughput Prediction in 5G Cellular Networks. In *2024 International Conference on Smart Applications, Communications and Networking (SmartNets)*. Institute of Electrical and Electronics Engineers (IEEE), 2024.

Boutiba, K., Bagaa, M., and Ksentini, A. Radio link failure prediction in 5G networks. In *2021 IEEE Global Communications Conference (GLOBECOM)*. IEEE, pp. 1–6, 2021.

Elsherbiny, H., Abbas, H. M., Abou-zeid, H., Hassanein, H. S., and Noureldin, A. 4g lte network throughput modelling and prediction. In *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*. IEEE, pp. 1–6, 2020.

Ghosh, A., Maeder, A., Baker, M., and Chandramouli, D. 5G Evolution: A View on 5G Cellular Technology beyond 3GPP Release 15. *IEEE Access* vol. 7, pp. 127639–127651, 2019.

Hewamalage, H., Ackermann, K., and Bergmeir, C. Forecast evaluation for data scientists: common pitfalls and best practices. *Data Mining and Knowledge Discovery* vol. 37, pp. 788–832, 3, 2023.

Hyndman, R. J. A brief history of forecasting competitions. *International Journal of Forecasting* vol. 36, pp. 7–14, 1, 2020.

Hyndman, R. J. and Athanasopoulos, G. *Forecasting: Principles and Practice.* OTexts, Melbourne, Australia, 2021.

Januschowski, T., Gasthaus, J., Wang, Y., Salinas, D., Flunkert, V., Bohlke-Schneider, M., and Callot, L. Criteria for classifying forecasting methods. *International Journal of Forecasting* vol. 36, pp. 167–177, 1, 2020.

Makridakis, S., Spiliotis, E., and Assimakopoulos, V. M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting* vol. 38, pp. 1346–1364, 10, 2022.

Narayanan, A., Ramadan, E., Mehta, R., Hu, X., Liu, Q., Fezeu, R. A. K., Dayalan, U. K., Verma, S., Ji, P., Li, T., Qian, F., and Zhang, Z.-L. Lumos5G: Mapping and Predicting Commercial mmWave 5G Throughput. In *Proceedings of the ACM SIGCOMM Internet Measurement Conference (IMC '20)*. Association for Computing Machinery (ACM), pp. 176–193, 2020.

Raca, D., Leahy, D., Sreenan, C. J., and Quinlan, J. J. Beyond throughput, the next generation: A 5G dataset with channel and context metrics. *MMSys 2020 - Proceedings of the 2020 Multimedia Systems Conference*, 5, 2020.

Santos, G. L., Endo, P. T., Sadok, D., and Kelner, J. When 5g meets deep learning: A systematic review. *Algorithms* vol. 13, pp. 208, aug, 2020.

Sharma, A., Pandit, S., and Talluri, S. R. Throughput prediction of fifth-generation cellular system using hybrid feature selection and enhanced sequential decision tree machine learning algorithm. *Wireless Networks* vol. 31, pp. 3025–3042, 2025.

Yeaser, K. M. A. and Hassan, K. M. A. 5G NR V2X Throughput Prediction Using Deep Hybrid Learning. In *Innovations in Electrical and Electronics Engineering: Proceedings of the 5th ICIEEL 2024*, A. Kalam, S. Mekhilef, and S. S. Williamson (Eds.). Springer Nature Singapore, Singapore, pp. 685–693, 2025.

Yingjie, Z. and Abolghasemi, M. Local vs. global models for hierarchical forecasting. , 11, 2024.