This article has been accepted for publication in IEEE Networking Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LNET.2023.3242680

Short-Term Multivariate KPI Forecasting in Rural Fixed Wireless LTE Networks

Alexander G. B. Colpitts, *Student Member, IEEE,* Brent R. Petersen, *Senior Member, IEEE*

***Abstract*—Time series forecasting has gained significant traction in LTE networks as a way to enable dynamic resource allocation, upgrade planning, and anomaly detection. This work investigates short-term key performance indicator (KPI) forecasting for rural fixed wireless LTE networks. We show that rural fixed wireless LTE KPIs have shorter temporal dependencies compared to urban mobile networks. Second, we identify that the inclusion of environmental exogenous features yields minimal accuracy improvements. Finally, we find that sequence-to-sequence-based (Seq2Seq) models outperform simpler recurrent neural network (RNN) models, such as long short-term memory (LSTM) and gated recurrent unit (GRU), and random forest (RF).**

***Index Terms*—Forecasting, Rural areas, Neural network application, Communication system performance**

I. Introduction

FORECASTS of performance-metrics for wireless networks are powerful for enabling intelligent networks. Activities like advanced fault detection and dynamic resource allocation can be leveraged to improve performance and observability [1]. In this work, we introduce and optimize short-term multivariate key performance indicator (KPI) forecasting in rural fixed wireless LTE networks using deep learning (DL). Forecasting in urban mobile networks has shown promising results [2]; however, forecasting KPIs for rural fixed wireless networks remains largely ignored despite the fact that rural environments are known to differ from urban environments in terms of both propagation characteristics [3] and user behaviour [4].

The application of DL to time series forecasting has received significant attention in recent years. Wireless communication networks have used DL with promising results for network management and monitoring [1]. DL and neural networks (NN) have become widely used because they reduce the burden of creating accurate statistical or empirical models of the processes that are forecasted. Many different NN models have been considered for time series forecasting. RNNs have shown remarkable success in time series forecasting [5]. Two variants of RNN have become popular for time series forecasting: long short-term memory (LSTM) and gated recurrent unit (GRU). LSTMs are an improvement over RNNs in that they introduce input, forget, and output gates to help address vanishing and exploding gradients to which vanilla RNNs are susceptible. Similarly, GRU uses update and reset gates to control the flow of information between cells; however, they do not contain internal memory and output their hidden state directly [5]. Both LSTM and GRU excel at learning long- and short-term dependencies of time series data.

The selection of a NN architecture depends on the application. For time series forecasting, some considerations are the number of time steps to predict, the relevant available features, and the statistical properties of the data. A survey of LSTM network architectures and their application to time series forecasting can be found in [6] while a practical overview of time series forecasting architectures can be found in [5]. The performance of different NN architectures in the context of LTE KPI forecasting has received little attention in the existing literature.

A large portion of wireless network forecasting research is focused on urban mobile LTE networks. Many studies have used data provided by the Telecom Italia Big Data Challenge. Although this data set has served as a sort of de facto standard for wireless network forecasting, the use of different data sets is critical to understanding model performance and architectural selection [7]. Existing forecasting research at the cell or network level usually focuses on cell load and wireless channel quality [2], [8].

The forecasting models used for LTE performance metrics have evolved as the fundamental understanding of time series forecasting models has improved. Before the rise of deep learning, statistical models such as ARIMA (autoregressive integrated moving average) and SARIMA (seasonal-ARIMA) were extremely popular. [9] utilized SARIMA and three weeks of data to estimate the cell load for the following week. A drawback of ARIMA and SARIMA is that they do not easily allow for the inclusion of exogenous variables in the model. Special events, holidays, and weather are known to alter user behavior and may further increase the accuracy of a forecasting model. It is also best practice with statistical models to retrain after every prediction, which significantly increases the computational cost of deploying a model.

A significant advantage of deep learning over traditional statistical methods is that the inclusion of exogenous features is straightforward. [2], [8], [10] utilize deep learning techniques for LTE load or channel quality forecasting. [10] compared several different regression techniques, including NNs, to predict the load of a cell. It was found that by adding spatio-temporal features to the model, the coefficient of determination, R^2, typically increased. Similarly, [8] leveraged spatial information for model augmentation but also included weather, special events, and network configurations into the model. The proposed model was more complex and used unique LSTMs to consider seasonality, periodicity, and locality, which were then fused with exogenous variables. [2] utilized a simple

Manuscript received ; revised . This work was supported by a MITACS Accelerate grant and Xplore Inc. (Corresponding author: Alexander G. B. Colpitts)

The authors are with the Department of Electrical and Computer Engineering, University of New Brunswick, Fredericton, NB, Canada (email: alex.colpitts@ieee.org; b.petersen@ieee.org).

This article has been accepted for publication in IEEE Networking Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LNET.2023.3242680

LSTM model with exogenous variables similar to those used in [8]. The results of [2], [8] concluded that the regression techniques based on NN outperformed naive forecasting, statistical methods, and tree-based regression techniques. It should be noted that all studies exclusively considered urban mobile wireless networks.

In this paper, we investigate deep learning for time series regression in a rural fixed-wireless network. For forecasting, we consider uplink and downlink throughputs and physical resource block (PRB) use, as well as the number of connected users. RNN-based forecasting models were considered, namely LSTM, GRU, and sequence-to-sequence-based models (Seq2Seq). In addition to comparing the performance of the aforementioned NN architectures, we demonstrate that the temporal dependencies of rural fixed wireless networks are much shorter than those of urban mobile networks considered in [9]. Furthermore, we identify that the inclusion of exogenous environmental features related to temperature and snow provides minimal benefit. This is noteworthy since urban mobile networks found a significant performance improvement by adding similar exogenous features [2]. Lastly, we identify that Seq2Seq-based models outperform both simpler RNN models, such as LSTM and GRU, and Random Forest regressors.

The rest of this paper is organized as follows: Section II discusses data sourcing, preliminary analysis, and the models to be used, Section III discusses the forecasting and feature selection results, and finally Section IV presents our concluding remarks.

II. Data, Procedure, and Analysis

One of the most powerful benefits of DL-based time series forecasting models is that the inclusion of exogenous features is straightforward. Although it is a simple task to add exogenous features, it is critical to ensure that they improve the model's performance. The selected endogenous LTE KPIs were chosen to help describe the overall user experience at the cellular level, namely uplink and downlink throughput and PRB use, as well as the number of connected users. To augment our data set, we consider three sets of exogenous features: temporal, network, and environment information. In this section, we outline the data sources, perform statistical analysis, and discuss the models used for forecasting.

A. Network Data

LTE network KPIs were collected from a fixed wireless rural network. Collection is typically done on a user, cell, or site basis. For this study, we consider data on a cellular basis. RAN KPI data was collected from 200 cells spread throughout New Brunswick, Canada. This region was selected because it would allow cells to be restricted to a single, broadly classified ecozone [11]. Data was collected from November 1, 2021 to March 31, 2022 with a sampling interval of 15 minutes. This period of time was selected to allow the inclusion of snow in the exogenous environmental features.

Five network KPIs were selected as endogenous variables because they concisely describe cell performance: number of active users, use of downlink PRB (DL-PRB), downlink cell throughput (DL-THRP), use of uplink physical resource blocks (UL-PRB) and uplink cell throughput (UL-THRP). In addition to the endogenous set of KPIs, there was the option to include other network KPIs as exogenous variables. The exogenous network KPIs were related to the uplink and downlink modulation and coding scheme (MCS).

B. Environment Data

Environment data was collected from the Historic Weather API of Environment Canada. In the region of interest, there are more than 200 weather stations. Each station provides at least one sampling interval (daily, hourly, monthly). For each of the sampling intervals, different features relating to the environment may be available. This work considers two environment features: snow depth and temperature. Temperature is available throughout all sampling intervals, whereas snow depth is only available on daily or larger sampling intervals. The network data is sampled on a 15-minute interval. To combine the environment features with the network data, we forward fill the environment data; in other words, we assume that the snow depth and temperature have not changed until the next reported environment data point. For temperature, this results in hourly samples being repeated 4 times, while snow depth is repeated 96 times.

C. Temporal Data

The structure of the time series data provides implicit temporal knowledge to the model, but does not explicitly provide the time of each input feature. Explicit temporal information was added by sinusoidally encoding the time values. Minutes, hours, days of week, days of month, and month of year were considered. To ensure that the temporal data were free from phase ambiguities, we encoded time values as sines and cosines.

D. Statistical Analysis

Time series forecasting is no exception to the adage *"Garbage in, garbage out."* Not all time series data sets are suitable for forecasting, making exploratory analysis critical to confirm whether it is worthwhile to proceed with modeling. An excellent statistical analysis of the KPIs of urban wireless networks can be found in [9]. The statistical analysis of this work is broken into two parts: stationarity testing and correlation testing. LTE cells and KPIs were independently analyzed, followed by meta-analysis to help understand the overarching behavior. The stationarity tests used a combination of the augmented Dickey-Fuller (ADF) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests. Correlation tests involved a complete and partial autocorrelation analysis to understand the periodicity of the data.

ADF and KPSS tests are used together to test stationarity and type of stationarity. The null hypothesis of the ADF is that the series has a unit root, while the alternate hypothesis is that it does not. The null hypothesis of the KPSS is that the series is trend-stationary, whereas the alternate hypothesis is that it has a unit root. The two tests allow for four outcomes: both tests

This article has been accepted for publication in IEEE Networking Letters. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/LNET.2023.3242680

indicate that it is not stationary, both tests indicate stationary, KPSS indicates that it is stationary while ADF does not, and ADF indicates that it is stationary while KPSS does not. If the tests do not agree, it indicates that the time series is trend- or difference-stationary, respectively. Most of the cells and associated KPIs were found to be stationary; in other words, ADF and KPSS both indicated that the data were stationary.

Understanding the periodicity of the data allows for an improved design of the forecasting model. Partial autocorrelation provides insight into the lag relationship between features. Passing a long time history to the model can increase accuracy if there is a relationship between the lags and the forecasting target. If the relationship is minimal for many of the lags, increasing the time history window will only serve to increase the computation time without increasing the accuracy. Partial autocorrelation analysis on a cell-by-cell basis found that most of the cells had short time dependencies. For a short-term forecast of 15 minutes, most cells only had strong correlations between the target and five time lags. Some cells had KPIs showing dependencies beyond five lags, so we considered a larger historical time window. Higher-lag relationships were significantly weaker than short-term lags, but there were some occurrences of moderate correlation up to 8 lags. Although some KPIs have longer temporal dependencies, it should be noted that the longer-term correlations are still relatively small and near zero in many cases. Partial autocorrelation values beyond the first five lags were often less than 0.1, while the first few lags had relationships much higher than 0.25. The autocorrelation and partial autocorrelation of DL-PRB for a single cell can be seen in Fig. 1. The temporal dependence of KPIs in rural fixed wireless networks has been found to be significantly different from their urban counterparts. The rural network considered exhibited minimal temporal dependencies beyond 2 hours, while urban networks exhibit significant relationships that last several weeks [9].

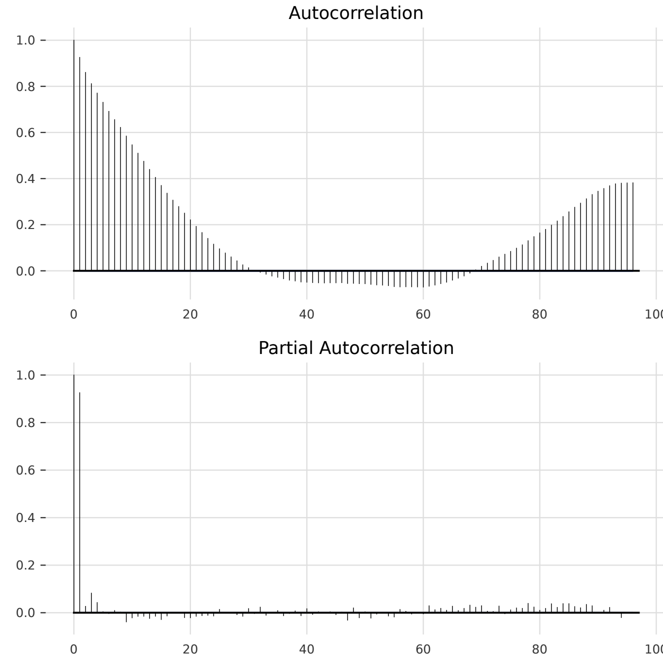

Figure 1: Autocorrelation and partial autocorrelation for DL-PRB for a single cell

E. Data Preparation

Preparation of data is the key for any forecasting problem, especially those using DL. The studies described in Section I that used NNs scaled their data to 0 to 1, also known as Min-Max scaling. Min-Max and standard scaling (zero mean, unit variance) are common forms of data scaling used before prediction. This work considers both global Min-Max and standard scaling, global scaling indicating that we scale the data based on all cells rather than on a cell-by-cell basis. Global scaling was used because the objective was to create a global model.

In addition to scaling, there is also the question of how to handle missing data. [8] utilized linear interpolation to handle missing time steps in the data. Another common method is to remove the missing time steps from the data set, as was done in [12]. We use zero-filling to accommodate missing time steps.

F. Models and Training

The objective of the models created in this work is to predict future KPI values based on historical input. Regular LSTM and GRU models have excellent long-term dependency learning capabilities but are typically outperformed by Seq2Seq models. [13] found that for short-term electricity consumption forecasting, the GRU-based Seq2Seq models outperformed the LSTM and vanilla RNN-based Seq2Seq models. Since the time horizon of the proposed forecasting task is short-term, we opt to use GRU-based Seq2Seq models with and without attention, which we will compare with LSTM and GRU models.

RNNs operate by feeding the data into the network one time step at a time. The network will learn from current and past time steps to make a future prediction. Although this does provide a way to learn time dependencies, vanilla RNNs are highly susceptible to exploding or vanishing gradients. LSTM networks have been proposed to overcome these problems. LSTM cells contain learned input and output control gates to manage the flow of information, thus protecting themselves from vanishing or exploding gradients [14]. GRU cells are similar to LSTM cells in that they manage the incoming flow of information into the cell; however, they differ in architecture in that GRU cells expose the hidden state directly to the output.

Seq2seq models leverage two unique networks in an encoder-decoder architecture. In this work, we use GRU-based Seq2Seq models. The input sequence is fed into the encoder network, which builds a compressed representation, often referred to as the context vector. The context vector is then decoded to obtain the target sequence [15]. Although they can learn more complex representations than RNN alone, Seq2Seq models can be limited by the amount of information that can be contained in the context vector. Attention has been proposed as a solution to this because it allows the model to attend to all the previous hidden states of the encoder [16]. When considering Seq2Seq models with attention, we use the scaled dot product attention as proposed in [17].

The models were trained for all combinations of including and excluding exogenous network, temporal, and environment features. A random forest (RF) regressor was used as a baseline model. This baseline was selected because tree-based regressors are known to achieve excellent performance in time series forecasting tasks [18]. Furthermore, RFs were used as a baseline model in [8].

For each combination of exogenous features, several sets of model hyperparameters were tested. Table I contains an overview of the selected hyperparameters. The data set for each model was made up of data from November 2021 to March 2022. Training and validation sets were created using a temporal slice across all cells. 80% of the data was used to train and the remaining 20% was used for validation. The data was temporally sliced to ensure that leakage between datasets did not occur. The maximum number of training epochs was 200 with a stopping criterion if the change in validation loss did not decrease. A patience value of 2 was used in the exit condition and the best weights were restored at the end of the training.

Table I: Hyperparameter options for model testing

Hyperparameter	Values
Number of encoder/decoder layers	[1, 2, 3]
Number of encoder/decoder cells per layer	[4, 8, 16, 32, 64]
Learning rate	1E-3
Batch size	128
Cost function	Mean-squared error (MSE)
Include environment (E)	[True, False]
Include time (T)	[True, False]
Include exogenous network (N)	[True, False]
History step size	8 (2 hours)
Data scaling	[Standard, Min-Max]

III. Results and Discussion

In this section, we compare the accuracy of various models with and without the use of exogenous features. Three sets of exogenous variables are considered: network, sinusoidal encoded times, and environmental data. The results presented indicate that the inclusion of exogenous variables for short-term multivariate forecasting provided minimal improvement in terms of RMSE (root mean squared error) or R^2.

First, we consider the models that perform best by model type. The baseline RF model achieved an RMSE of 5.013 and an R^2 of 0.858. As seen in Table II, all the best models by type outperformed the baseline where the Seq2Seq-based model performed the best. The inclusion of an attention mechanism in the Seq2Seq models was found to have a small negative impact on the accuracy of the forecast. Although the top models shown in Table III have 64 hidden units with predominantly three layers, many models with 32 hidden units achieve similar results. This result aligns with the findings of [19] in terms of model sizes and contrasts with the hyperparameter choices of many existing works that used many more layers or hidden units [2], [20], [21]. A summary of the best models of each type can be found in Table II. A box plot of the training results for the top 40 of each type of model can be seen in Fig. 2. It can be observed that Seq2Seq with and without attention achieved a similar RMSE, while LSTM and GRU had a higher RMSE and had a higher variance in the top models. Moreover, we found that all of the best performing models leveraged standardization over Min-Max scaling, which contrasts with the use of Min-Max scaling in other works.

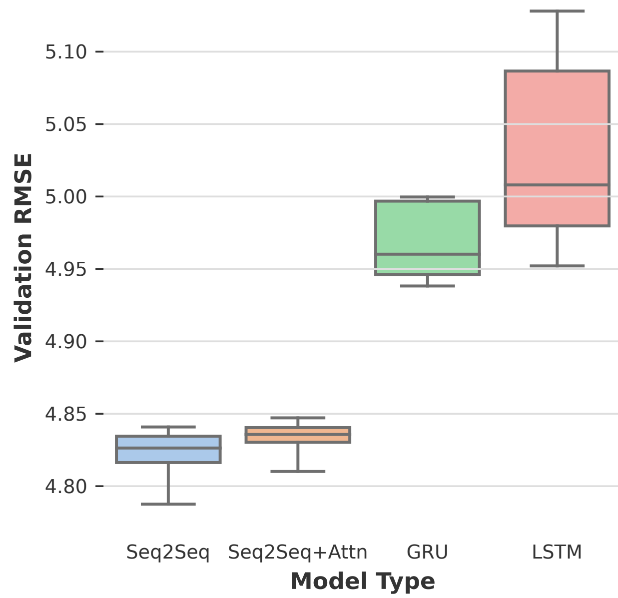

Figure 2: Box plot of training results for the top 40 models for each of the four types of models tested.

Table II: Best results by type of model. N, T, and E represent the inclusion of network, time, and environment exogenous features, respectively.

RMSE	R^2	N	T	E	Hidden	Layers	Type
4.788	0.868	False	True	False	64	2	Seq2Seq
4.810	0.867	True	True	True	64	2	Seq2Seq+Attn
4.938	0.862	True	False	True	32	1	GRU
4.952	0.862	True	False	False	32	1	LSTM
5.013	0.858	False	False	False	-	-	RF

The best-performing model for each combination of hyperparameters can be found in Table III. It should be noted that the models that performed the best in all combinations of features were Seq2Seq, except for the case where only environment features were added. Seq2Seq with and without attention dominated the top performing spots; however, Seq2Seq without attention achieved the best results for all but one combination. Furthermore, the performance difference between the best model for each hyperparameter combination was found to be approximately 1%, indicating that inclusion had minimal benefit.

We have identified that the inclusion of exogenous variables was found to have a minimal effect on single-step forecasting accuracy; however, we found that the size of the model also had a minimal impact on the accuracy. This result comes as no surprise after the statistical analysis in Section II-D demonstrated that the KPIs exhibit minimal temporal dependencies beyond

This article has been accepted for publication in IEEE Networking Letters. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/LNET.2023.3242680

Table III: Best results for each combination of exogenous hyperparameters. N, T, and E represent the inclusion of network, time, and environment exogenous features, respectively.

RMSE	R^2	N	T	E	Hidden	Layers
4.788	0.868	False	True	False	64.000	2
4.807	0.867	True	True	True	64.000	3
4.808	0.868	True	False	True	64.000	2
4.811	0.868	True	True	False	64.000	3
4.813	0.868	False	True	True	64.000	1
4.819	0.867	True	False	False	64.000	3
4.834	0.867	False	False	False	64.000	3
4.837	0.867	False	False	True	64.000	1

a few time steps. The smaller models had marginally lower accuracy than their larger counterparts, but the computational requirements were significantly different. The smallest model tested contained 100s of trainable parameters, while the largest contained more than 100,000. The best performing models had between 10,000-140,000 parameters. The size of the model becomes a major concern for some forecasting applications. If network KPI forecasting cannot be completed within the window of interest, the results cannot be used for tasks such as dynamic resource allocation or real-time anomaly detection.

Our findings have indicated that exogenous temporal, network, and environmental features cannot be leveraged to significantly improve short-term forecasting in rural fixed wireless LTE networks. These findings come with caveats, such as that the environment data is only available with a sampling period of 60 minutes and that temporal information can be implied from the structure of the data. This is a notable difference from existing urban studies.

IV. Conclusion

In this work, we present the results of the selection of exogenous features for the prediction of short-term KPI in a rural fixed wireless LTE network. Statistical time series analysis found that the temporal dependencies for a single KPI were very short, often fewer than four time steps (one hour), which is dramatically shorter than those reported for urban networks. The data set was enhanced using additional network KPIs, sinusoidal encoded time features, and exogenous environment features, namely temperature and snow depth. We found that the inclusion of exogenous variables did not produce a significant improvement in the models considered. This is a notable difference from the results found in similar urban mobile network studies, which found that exogenous features significantly improved the accuracy of the forecast.

Although we have identified that Seq2Seq models with and without attention outperform LSTM, GRU, and RFs, the use of different model architectures such as temporal convolutional NN, graph NN, and transformer-based models could achieve better performance. Moving forward, a survey of different network architectures could be performed to greatly improve the understanding and accuracy of KPI forecasting at the cellular level in rural fixed wireless networks.

References

[1] C. Zhang, P. Patras, and H. Haddadi, "Deep Learning in Mobile and Wireless Networking: A Survey," *IEEE Commun. Surv. Tutorials*, vol. 21, no. 3, pp. 2224–2287, 2019.

[2] T. Kuber, I. Seskar, and N. Mandayam, "Traffic prediction by augmenting cellular data with non-cellular attributes," *IEEE Wirel. Commun. Netw. Conf. WCNC*, vol. 2021-March, pp. 1–6, 2021.

[3] V. S. Abhayawardhana, I. J. Wassellt, D. Crosby, M. P. Sellars, and M. G. Brown, "Comparison of empirical propagation path loss models for fixed wireless access systems," *IEEE Veh. Technol. Conf.*, vol. 61, no. 1, pp. 73–77, 2005.

[4] A. L. Wagner, F. Keusch, T. Yan, and P. J. Clarke, "The impact of weather on summer and winter exercise behaviors," *J. Sport Heal. Sci.*, vol. 8, no. 1, pp. 39–45, 2019. [Online]. Available: https://doi.org/10.1016/j.jshs.2016.07.007

[5] P. Lara-Benítez, M. Carranza-García, and J. C. Riquelme, "An Experimental Review on Deep Learning Architectures for Time Series Forecasting," *Int. J. Neural Syst.*, vol. 31, no. 3, pp. 1–25, 2021.

[6] B. Lindemann, T. Müller, H. Vietz, N. Jazdi, and M. Weyrich, "A survey on long short-term memory networks for time series prediction," *Procedia CIRP*, vol. 99, pp. 650–655, 2021. [Online]. Available: https://doi.org/10.1016/j.procir.2021.03.088

[7] S. Jung, K.-M. Kim, H. Kwak, and Y.-J. Park, "A Worrying Analysis of Probabilistic Time-series Models for Sales Forecasting," *arXiv*, pp. 1–11, 2020. [Online]. Available: http://arxiv.org/abs/2011.10715

[8] A. Bhorkar, K. Zhang, and J. Wang, "DeepAuto: A Hierarchical Deep Learning Framework for Real-Time Prediction in Cellular Networks," *arXiv*, 2019. [Online]. Available: http://arxiv.org/abs/2001.01553

[9] F. Xu, Y. Lin, J. Huang, D. Wu, H. Shi, J. Song, and Y. Li, "Big Data Driven Mobile Traffic Understanding and Forecasting: A Time Series Approach," *IEEE Trans. Serv. Comput.*, vol. 9, no. 5, pp. 796–805, 2016.

[10] E. Lovisotto, E. Vianello, D. Cazzaro, M. Polese, F. Chiariotti, D. Zucchetto, A. Zanella, and M. Zorzi, "Cell traffic prediction using joint spatio-temporal information," *2017 6th Int. Conf. Mod. Circuits Syst. Technol. MOCAST 2017*, 2017.

[11] Natural Resources Canada, "Statistical data — Forests — Natural Resources Canada," 2020. [Online]. Available: https://cfs.nrcan.gc.ca/statsprofile/overview/NB

[12] X. Wang, Z. Zhou, F. Xiao, K. Xing, Z. Yang, Y. Liu, and C. Peng, "Spatio-Temporal Analysis and Prediction of Cellular Traffic in Metropolis," *IEEE Trans. Mob. Comput.*, vol. 18, no. 9, pp. 2190–2202, 2018.

[13] L. Sehovac, C. Nesen, and K. Grolinger, "Forecasting building energy consumption with deep learning: A sequence to sequence approach," *Proc. - 2019 IEEE Int. Congr. Internet Things, ICIOT 2019 - Part 2019 IEEE World Congr. Serv.*, pp. 108–116, 2019.

[14] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training Recurrent Neural Networks," *arXiv*, 2013.

[15] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Adv. Neural Inf. Process. Syst.*, vol. 4, no. January, 2014, pp. 3104–3112.

[16] M. T. Luong, H. Pham, and C. D. Manning, "Effective approaches to attention-based neural machine translation," *Conf. Proc. - EMNLP 2015 Conf. Empir. Methods Nat. Lang. Process.*, pp. 1412–1421, 2015.

[17] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Adv. Neural Inf. Process. Syst.*, vol. 2017-Decem, no. Nips, pp. 5999–6009, 2017.

[18] R. P. Masini, M. C. Medeiros, and E. F. Mendes, "Machine Learning Advances for Time Series Forecasting," in *J. Econ. Surv.*, no. Ml, 2021, pp. 1–36.

[19] J. Feng, X. Chen, R. Gao, M. Zeng, and Y. Li, "DeepTP: An End-to-End Neural Network for Mobile Cellular Traffic Prediction," *IEEE Netw.*, vol. 32, no. 6, pp. 108–115, 2018.

[20] F. W. Alsaade and M. Hmoud Al-Adhaileh, "Cellular Traffic Prediction Based on an Intelligent Model," *Mob. Inf. Syst.*, vol. 2021, 2021.

[21] H. S. Jang, H. Lee, H. Kwon, and S. Park, "Deep Learning-Based Prediction of Resource Block Usage Rate for Spectrum Saturation Diagnosis," *IEEE Access*, pp. 59 703–59 714, 2021.