60th International Scientific Conference on Information, Communication and Energy Systems and Technologies,
Ohrid, North Macedonia, June 26–28, 2025

Time Series Prediction of Electricity Consumption in North Macedonia Based on LSTM and Feature Analysis

Nikola Gachevski[1], Mitko Kostov[1], Metodija Atanasovski[1]

Abstract – **In this paper, a model for hourly electricity load forecasting for the Republic of North Macedonia has been presented using deep learning. A soft attention-based LSTM network was trained on the period 2016-2020 and later validated through the usage of the consumption for 2021. It uses meteorological as well as temporal inputs and delivered a MAPE value of 4.93%. Feature importance analysis revealed that day type and temperature had the highest impact on prediction. The model is very accurate, with more deviations observed at national holidays and during changes of seasons.**

Keywords –**Load Forecasting, Machine Learning, LSTM, Attention Mechanism, Energy Consumption**

I. Introduction

The increasing complexity of modern energy systems has brought forth the necessity of advanced forecasting models that can accurately predict electricity demand. Accurate load forecasting enables energy suppliers to balance supply and demand, optimize operational efficiency, and minimize economic losses. Traditional statistical approaches, while sufficient in steady systems, do not capture the dynamic and nonlinear characteristics of modern patterns of power consumption, particularly during periods of instability such as global health pandemics or holiday-induced shifts in consumerism.

In the last couple of years, deep learning methods have gained a huge pace in time series prediction due to the fact that they are able to learn hierarchical data patterns automatically. Among them, Long Short-Term Memory (LSTM) networks, being a type of recurrent neural network (RNN), have emerged as highly promising in sequential relations and long-range temporal patterns. Their architecture enables the network to recall information from long input sequences, and thus they are most appropriate for hourly energy consumption data [1].

In this study, an LSTM-hybrid model with a soft attention mechanism is proposed to forecast hourly electricity demand in the Republic of North Macedonia for 2021. The model is trained on a five-year historical dataset (2016–2020) with both temporal and contextual features. The attention mechanism is employed to enhance the interpretability of the model by dynamically focusing on significant time steps, thereby improving overall performance.

Moreover, a gradients-based feature importance analysis was incorporated to evaluate the contribution of each input variable.

[1]Nikola Gachevski, Mitko Kostov and Metodija Atanasovski are with University St. Kliment Ohridski - Bitola, Faculty of Technical Sciences, Makedonska falanga 37, 7000 Bitola, Republic of North Macedonia, E-mail: {nikola.gacevski, mitko.kostov, metodija.atanasovski}@uklo.edu.mk.

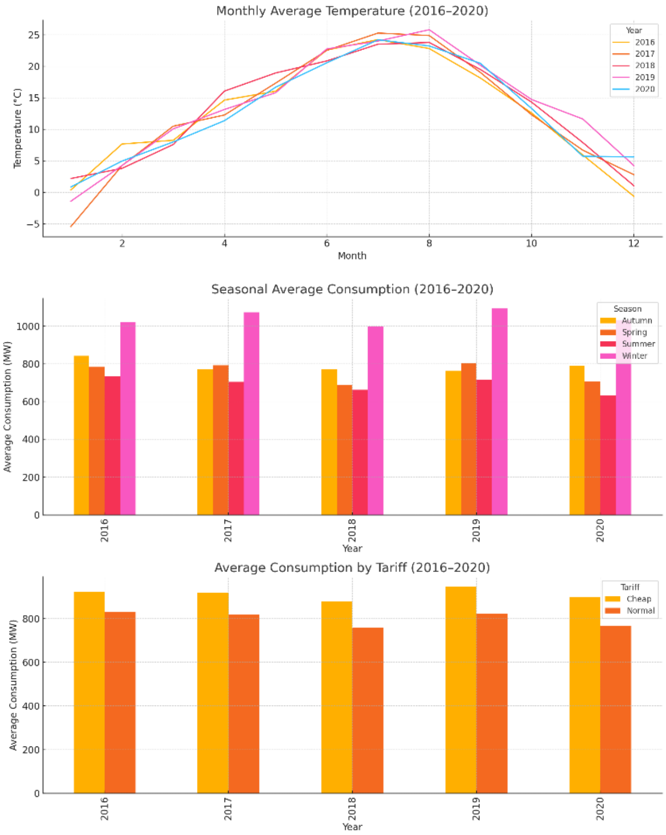

Fig. 1. Graphical breakdown of temperature profiles, seasonal load, and impact by tariff (for the years 2016–2020

The results indicated that temperature, whether it was a workday or non-workday, and off-peak tariff times were among the most determining factors of energy demand, while seasonal indicators and the pandemic variable had a lesser effect [2][3].

The last model generated a mean absolute percentage error (MAPE) of 4.93%, showing excellent predictive accuracy. Detailed examination of cases with higher forecasting deviation, like religious holidays and heatwaves, revealed clear patterns of behavior not always present in training data, offering valuable insight into model shortcomings and areas for future improvement.

The paper is organized as follows. After Introduction, the training and testing dataset for the model and data preparation techniques are presented in Section II. Section III presents the methodology, including the LSTM network structure with the attention mechanism and training procedure. Section IV gives the results and in-depth discussion on the performance of the model and impact of the significant features. Section V concludes the study and recommends future studies.

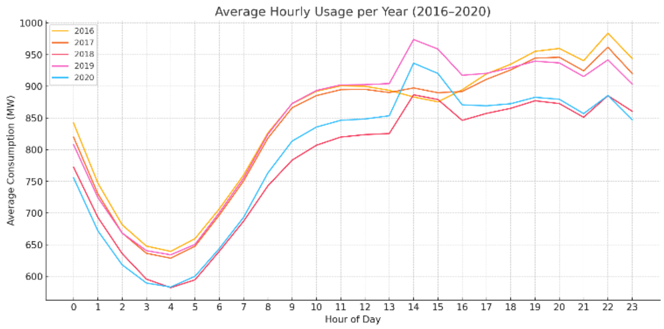

Fig. 2. Average hourly usage per year of the training set (2016–2020), showing diurnal and seasonal patterns.

II. Data Description

The analysis is based on a comprehensive dataset of hourly electricity consumption in the Republic of North Macedonia, covering the five-year period from 2016 to 2020 [4][5]. Historical data for 2016-2020 is used for training and only the year 2021 is used for testing and validation alone. Each observation in the data has not only the actual energy consumption but also some exogenous variables that are known or predictable in advance and are therefore suitable for forecasting.

One of the characteristics included is the ambient temperature, a key driver of consumption, especially in times of extreme weather [2][6]. The data also includes a flag for whether the day is a working or non-working day, season classification, and a binary flag for pandemic conditions, which is particularly relevant to the year 2020 due to the COVID-19 pandemic [7]. The CE (Cheap Energy) feature also captures low-tariff periods, which tend to coincide with normal patterns of consumer behavior and therefore are highly predictive. A time-feature transformation was applied by encoding the hour of day using sine and cosine transformations in an effort to better represent the cyclical patterns of electricity usage. The input numbers were all scaled by a Min-Max scaler to rescale values into the [0, 1] range in order to enable the model to train better [1].

The evolution and variability of these characteristics during the five-year period are visible in Figure 1, a graphical breakdown of temperature profiles, seasonal load, and impact by tariff. The average hourly usage per year of the training set is shown in Figure 2, demonstrating cyclic diurnal and seasonal patterns.

III. Methodology

The core of the predictive framework used here is a Long Short-Term Memory (LSTM) neural network [8], a specialized type of Recurrent Neural Network (RNN) [8][9] renowned for its ability to learn long-term dependencies in sequences. The LSTM architecture is well suited to learning temporal data such as electricity consumption, where current values are highly dependent on previous time instances. The model learned to forecast electrical demand hourly in the year 2021 based on patterns learned from historical data from 2016 to 2020.

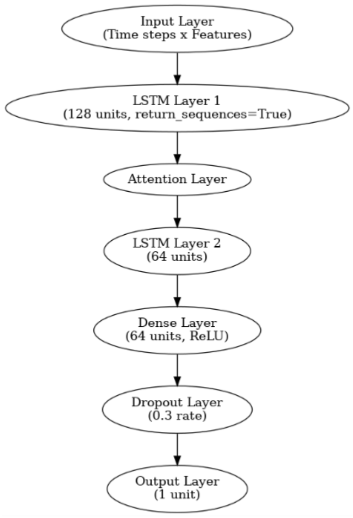

Fig. 3. Model architecture schematic, including the attention mechanism.

A soft attention mechanism was also added in order to further enhance model performance and interpretability. This layer imposes adaptive weights over the LSTM's hidden states across all time steps, effectively allowing the model to "attend" to the most salient inputs in the sequence. This attention mechanism augmented the model's ability to detect and order relevant patterns, particularly under volatile consumption periods [10]. A schematic diagram of the whole architecture, including the attention module, is depicted in Figure 3.

The network architecture consists of an input layer and two LSTM layers with 128 and 64 hidden units, respectively. After the use of the attention mechanism, the dense fully connected layer with the ReLU activation is used and then a dropout layer to prevent overfitting [11][12]. Finally, a single-node output layer generates the predicted value of consumption. The model is compiled using the Mean Squared Error (MSE) loss function and Adam optimizer for optimizing. The Adam algorithm was chosen due to its adaptive learning rate properties and strong performance on time series forecasting tasks. In addition, a learning rate scheduler with exponential decay was used to reduce the learning rate slowly during training and improve convergence stability [12].

There was training for over 100 epochs with a batch size of 32 and a validation split of 20% to monitor the generalization capability of the model. Input data preparation was done using a sliding window technique, where every input sequence had 48 time steps (two days) to forecast the next hour's consumption. This time window was chosen after experimental testing to strike a balance between model complexity and forecasting accuracy [13].

For comparing the contribution value made by all input features after training, the Integrated Gradients (IG) approach was adopted. IG belongs to the group of gradient-based attribution methods where each feature is weighted by cumulating the gradients along a line path from the baseline input towards actual input. The results of the feature importance analysis are shown in Figure 4, where workday classification, CE (cheap energy) periods, and temperature

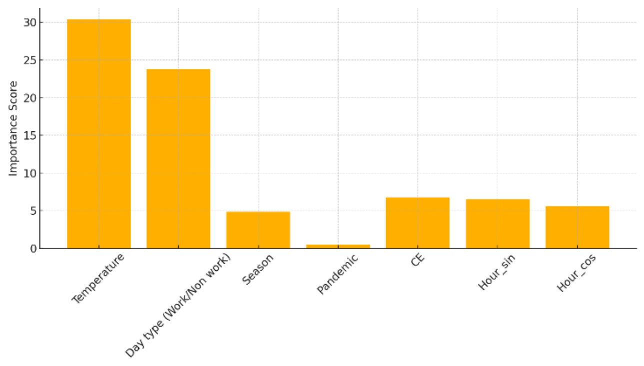

Fig. 4. Feature importance analysis using Integrated Gradients – bar chart showing importance of each feature.

were the major contributing factors to the prediction. These findings are then cross-checked against actual observed anomalies in consumption, further reinforcing the explanatory power of the selected features.

IV. Results and Discussion

The model was cross-validated on unseen data for the year 2021 using Mean Absolute Percentage Error (MAPE) as the performance metric [14]. The model's last MAPE was 4.93%, which reflects extremely high precision and reliability in forecasting electricity demand for an entire calendar year. Such accuracy is extremely competitive, particularly because of seasonality and fluctuation in power usage in North Macedonia.

A study from North Macedonia on day-ahead electricity load forecasting using ARIMA-based models found that a basic ARIMA model achieved a Mean Absolute Percentage Error (MAPE) of around 5%, while incorporating temperature as an exogenous input in a SARIMAX model reduced the error to approximately 3.6% [15]. These error rates are in the same range as the LSTM with attention (MAPE 4.93%), suggesting comparable accuracy between the classical statistical approach and the deep learning method within this regional context.

MAPE is an appropriate measure of load forecasting since it expresses error in prediction as a percentage of the actual demand, thus making the practitioner comprehend the accuracy with ease. It is scale-invariant – a 5% error is the same regardless of whether maximum demand is 500 MW or 1500 MW and standard practice to compare forecasting models in power systems [10]. By expressing errors in terms of percentage, MAPE allows easy comparison of model accuracy across different load levels (e.g., an MAPE of 4.93% means the model forecasts are ~4.93% different on average from actual). This makes MAPE a sound means of assessing accuracy of electricity demand forecasts in this research.

A feature importance analysis using the Integrated Gradients method ranked temperature as the most influential, followed by workday classification and tariff-related attributes. Pandemic and seasonality, while in the model, had zero contributions towards its predictive capacity. These findings are further illustrated through graphical analyses in Figure 5 and Figure 6 which show the relationship between actual and predicted consumption and the impact of prominent features.

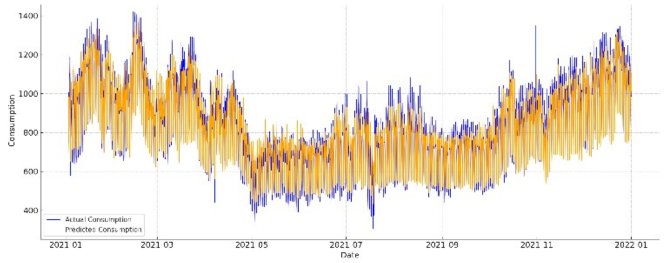

Fig. 5. Full-year 2021 comparison between actual and predicted consumption.

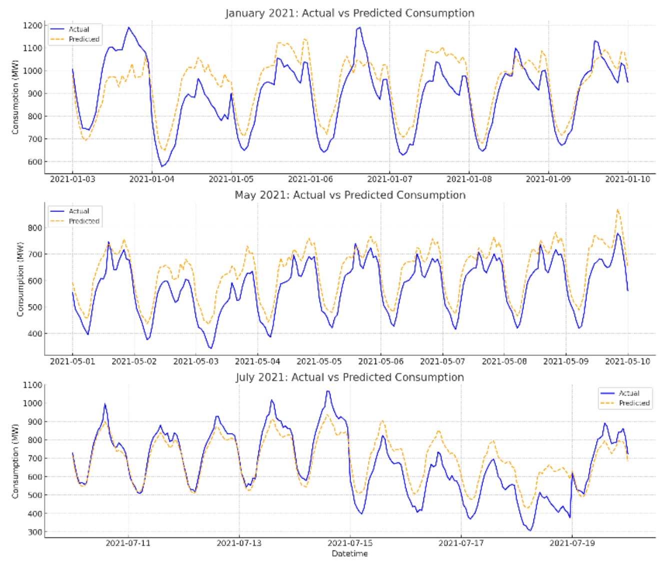

Fig. 6. Zoomed-in comparison of actual vs predicted values during periods with largest deviation (January, May, July).

Figure 5 is a comparative plot of predicted and actual hourly consumption throughout the year. While the model captures the overall trends and seasonal behavior, there are some phases with considerable deviations. A closer look, as presented in Figure 6, indicates that the largest prediction errors happen on specific dates in January, May, and July.

The most heavily weighted top ten highest average error predictive days were calculated which included January 3rd, 4th, and 7th, May 2nd–4th, and July 15th–18th. These are significant national and religious public holidays such as Christmas and Easter, and Eid al-Adha, when usual consumption patterns significantly diverge from previous patterns due to social festivities, mobility, and alterations in daily habits [16].

To determine the cause of the consumption anomalies, the temperature record of the five years from 2016 to 2020 was compared with that of 2021. The comparison revealed that the temperatures on the anomalous days in July 2021 were slightly above the five-year average. These elevated temperatures, coupled with social dynamics in the post-pandemic period—e.g., increased mobility of the population and reverse migration—were likely to be the cause of the unexpected increase in electricity demand [16]. These factors may not have been properly captured by the model, thus leading to deviations between expected and actual values.

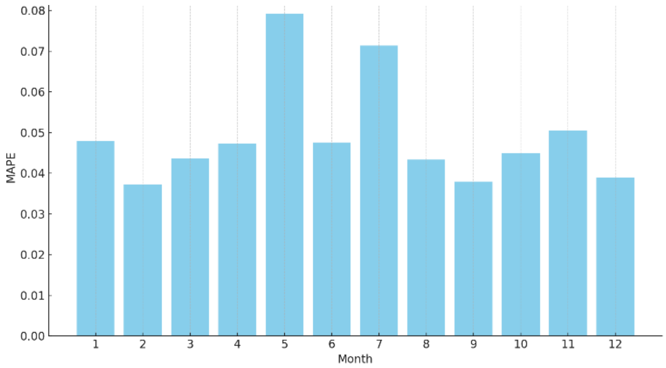

Fig. 7. Monthly error analysis – MAPE per month bar chart.

The use of an attention mechanism built into the LSTM model played a key role in strengthening the prediction for most of the year. Nevertheless, attention weights did fall off in the case of sudden holiday-related changes, suggesting that additional fine-tuning in the identification of such contextual variables may be required.

In addition, the Integrated Gradients-based feature importance analysis uncovered that the most influential features in model decision-making were Temperature, CE hours (low-cost energy hours), and Day Type (Work/Non-work). Of note, pandemic status and seasonal category were comparatively less impactful, indicating that real-world environmental input along with overt temporal structure (hour of day and day classification) provides more predictiveness.

Monthly error analysis presented in Figure 7 confirmed that the model performed best for months with solid and stable consumption patterns such as April and October and that more variances were observed in January, May, and July. This indicates the suitability of the application of behavioral and event-driven information in future models to continue minimizing prediction error for outlier months.

V. Conclusion

In this paper, an attention-based LSTM hybrid model was proposed and applied to predict hourly electricity demand in the Republic of North Macedonia in 2021. The model was fitted on five years of data from 2016 to 2020 based on a suitably selected feature set including temperature, type of day, season dummies, pandemic effect, and price time (CE hours). MAPE was used to measure the performance of the model, and its final value came out to be 4.93%, a very high prediction accuracy.

The careful analysis of prediction errors revealed the largest errors were during periods linked with national holidays and religious celebrations, and during heatwaves and season changes. This illustrates the issue of detecting erratic patterns in activity and bringing in additional context such as holiday calendars and social events into future modeling.

The incorporation of an attention mechanism improved the attention of the model to the most important temporal features, and analysis of feature importance confirmed the dominance of temperature and time-based measures in consumption patterns. Performance aside, the research further illustrates that greater gains can still be obtained through better temporal context modeling, especially during periods of socio-cultural importance.

Overall, the proposed LSTM-based system was a legitimate and comprehensible energy demand forecasting model. The model will, in future research, be enhanced and generalized further by incorporating additional external data sources such as calendar-dependent public activities, mobility routines, and home behavioral factors to tackle heterogeneous consumption scenarios.

References

[1] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory.

[2] Popovski, P., Kostov, M., Atanasovski, M. and Veljanovski G (2020). Power Load Forecast for North Macedonia Using Machine Learning.

[3] Ministry of Economy, Republic of North Macedonia. Energy Strategy 2040. Government Publication, 2021.

[4] https://www.mepso.com.mk/.

[5] Atanasovski M., Kostov M., Arapinoski B., and Spirovski M., (2020), K-Nearest Neighbor Regression for Forecasting Electricity Demand.

[6] https://www.wunderground.com/history.

[7] Veljanovski G., Atanasovski M., Kostov M and Popovski P. Application of Neural Networks for Short Term Load forecasting in Power System of North Macedonia (2020).

[8] Sherstinsky, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network.

[9] Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). A Critical Review of Recurrent Neural Networks for Sequence Learning.

[10] Qin, Y., Song, D., Chen, H., Cheng, W., Jiang, G., & Cottrell, G. (2017). A dual-stage attention-based recurrent neural network for time series prediction.

[11] Yuzhen L., Fathi M. S., Simplified Gating in Long Short-term Memory (LSTM) Recurrent Neural Networks (2017).

[12] Agarap A. F. M. Deep Learning using Rectified Linear Units (ReLU) (2018).

[13] Kingma P. D., Lei B. J., Adam: A method for stochastic optimization (2014).

[14] Bas E., Egrioglu E., Yolcu U., Bootstrapped Holt Method with Autoregressive Coefficients Based on Harmony Search Algorithm.

[15] Kotevska A., Rogleva K.N., Short-term load forecasting using time series analysis: A case study for the republic of North Macedonia (2020).

[16] https://www.mtsp.gov.mk/.