Daily Electricity Consumption Forecasting: A Comparative Study of Neural Network and Radial Basis Function Models

Agresa Qosja, Didier Georges, Eralda Gjika, Ligor Nikolla, Arben Cela

Daily Electricity Consumption Forecasting: A Comparative Study of Neural Network and Radial Basis Function Models

Agresa Qosja[1], Didier Georges[2], Eralda Gjika[3], Ligor Nikolla[4], Arben Cela[5]

***Abstract*— Electricity consumption forecasting stands as a critical research domain within electrical engineering, with myriad of traditional forecasting models and artificial intelligence techniques undergoing rigorous examination. This paper is devoted to a comparison of three Machine Learning approaches to surrogate modelling and forecasting of the daily electricity consumption in Tirana, Albania: A Radial Basis Function (RBF) approach, a feed forward Neural Network approach and a Recurrent Neural Network approach. Through meticulous experimentation across four distinct scenarios encompassing variations in training/testing splits, historical data utilization, and hyper parameter optimization, we thoroughly evaluate the performance of each model. Comparative analysis is conducted based on model fit, computational efficiency, and error measurement metrics. Our findings highlight the remarkable performance of the RBF ARX approach, underscoring its effectiveness in accurately forecasting electricity consumption.**

Key-words: Electricity consumption forecasting, neural network modelling, Radial Basis Function modelling.

I. INTRODUCTION

The ability to predict electricity demand is a key element in improving the management of power generation facilities, particularly in the context of the introduction of renewable energy sources and climate change. In light of this, there exists a necessity for novel applied models tailored to the nuanced circumstances of the Albanian context, forming the foundation of this study. Albania predominantly relies on renewable energy generation, ranging from 53.3% in the least favorable production scenario in 2022 to 71.3% in the optimal scenario of 2016 [1]. Applying an accurate model for the next day's consumption helps the respective institutions to determine the precise amount of energy the country needs to import from interconnected lines. As production is predominantly reliant on precipitation or reserved water, with 99% sourced from hydropower plants, consumption is influenced by temperature fluctuations typical of the Mediterranean climate region.

The prediction of electricity consumption has emerged as a prominent area for the application of innovative techniques stemming from recent advancements in neural network research. In their study, [2], the authors showcase the efficacy of deep learning models in short-term load forecasting using deep neural networks. The demonstrated accuracy of these models suggests potential for further applications in the field.

In a comprehensive analysis of the relationship between electricity consumption and economic activities [3], researchers investigated long- and short-run dynamics using panel data spanning 160 countries over a 30 year period. This study, which accounted for variables such as per capita GDP, degree of electricity dependence, and urbanization levels, revealed nuanced insights into the electricity-growth nexus across different regions and income level. The correlation between economic activity and electricity consumption, particularly influenced by household heating and cooling systems, is introduced by [4]. Authors used ANN models to predict over 12 regions in Turkey, emphasizing the importance of incorporating temperature as an exogenous variable.

In the line of considering the influence of exogenous variables on electricity consumption, researchers in [5] underscore the significant impact of weather on electricity demand, particularly within the context of Mediterranean climate change in Italy over a 15-year period. By employing various regression models that incorporate temperature, wind speed, relative humidity, and cloud cover, the study endeavors to forecast electricity demand up to one month ahead. The findings emphasize the superior accuracy of models utilizing temperature as a primary input compared to alternative models. Neural networks models are largely used in the field of electricity consumption forecasting for short- and long-term scenarios. From the analysis over [6] we have that Machine Learning models combined with the dimensionality reduction techniques have a better performance than other pure machine learning models.

Recurrent Neural Network (RNN) models can enhance prediction results by iteratively refining information in each row of the hidden layers, leveraging connections with the previous rows to capture temporal dependencies and improve forecasting accuracy [7]. From the applications of RNN models in electricity consumption, [8] provides a comparison and overview of these models in short-term load forecasting, suggesting that Long Short Term Memories (LSTM) and Gated Recurrent Units (GRU) layers exhibit superior performance in capturing highly nonlinear relationships between datasets. Also the results from [9] employ deep learning algorithms, including LSTM, GRU, and RNN, to forecast electrical loads based on current measurements, with the GRU model achieving the highest accuracy and lowest error. Furthermore, forecasting electricity consumption involves various models, as evidenced by research on an engineering

[1] Agresa Qosja, Univ. Grenoble Alpes, CNRS, Grenoble INP-UGA, GIPSA-lab, 38000 Grenoble, France, and University Metropolitan Tirana, Albania, agresa.qosja@grenoble-inp.fr

[2] Didier Georges, Univ. Grenoble Alpes, CNRS, Grenoble INP-UGA, GIPSA-lab, 38000 Grenoble, France, didier.georges@grenoble-inp.fr

[3] Eralda Gjika (Dhamo), Department of Applied Mathematics, University of Tirana, Albania, eralda.dhamo@fshn.edu.al

[4] Ligor Nikolla, University Metropolitan Tirana, Albania, ligor.nikolla@umt.edu.al

[5] Arben Cela, UGE-ESIEE Paris, France, arben.cela@esiee.fr

company in Russia [10]. This study covers RNN, SVM-RBF, SVM linear, and statistical models, emphasizing their relevance for analyzing hourly data within the period under consideration. The results from [11] compare RNN models with classical statistical approaches using monthly testing periods, showing that neural networks outperform the classical methods. This research focuses on identifying variables with the most significant impact on the prediction model. The evaluation of model performance is conducted across various monthly periods throughout the year. The applications of statistical models and artificial neural networks over a three-year dataset in the Albanian case, are presented in [12]. The findings indicate that hybrid neural networks outperform classical statistical models. While the the economical activity effect is reflected in the real dataset over the week days and week-end days. This is related with the feature of seasonality in variables.

This paper investigates the effectiveness of different machine learning methods for predicting day-ahead electricity consumption, utilizing historical consumption data and temperature as an exogenous variable, with a focus on a data-driven case study spanning 34 months in Tirana, Albania. Our study encompasses the use of state-of-the-art methodologies including a feed forward Neural Network (ANN), a Recurrent Neural Network (RNN) with simple layers or gated Recurrent Unit (GRU) layers, and Radial Basis Function Networks with a Gaussian kernel function (RBF). These models have played a pivotal role in enhancing the accuracy and efficiency of predicting energy consumption patterns, thereby facilitating optimized grid operation and energy management.

The paper is structured as follows: The Methodology section describes the Machine Learning techniques used in this paper, including a Radial Basis Function Network, a feed forward Neural Network, and a Recurrent Neural Network, for predicting energy consumption. The Dataset section provides an overview of the dataset features. The Results section presents the findings and discrepancies observed across different scenarios. This is followed by the Conclusions section, which outlines the conclusions drawn from the models' performance and offers suggestions for future work.

II. Methodology

This paper employs both neural networks and radial basis functions to forecast day-ahead electricity consumption while considering the average temperature as an exogenous variable within a daily dataset. Specifically, we explore the following models: A RBF network, A feed forward Neural Network, a Recurrent Neural Network with a simple structure, or incorporating Gated Recurrent Unit (GRU) layers. Across all models, we adopt a three-layer architecture comprising an input layer, a hidden layer, and an output layer. These models are designed to capture the complex relationships between variables and generate accurate predictions of electricity consumption.

Our main objective is to forecast energy consumption at a given day t, $C(t)$, on the basis of a nonlinear auto regressive (ARX) model, defined as a function of both historical energy consumption and the impact of temperature variations in the city. The general prediction function is represented as follows:

$$\begin{aligned} C(t) = F\big[&C(t-1), C(t-2), \ldots, C(t-N_C), \\ &T(t-1), T(t-2), \ldots, T(t-N_T)\big] \end{aligned} \tag{1}$$

where

- C - is the consumption variable,
- T - is the exogenous variable, in this case, temperature,
- t - is the daily index time
- N_C - is the number of past daily samples of the energy consumption
- N_T - is the number of past daily samples of the temperature

A. RBF Approach

RBFs have demonstrated exceptional effectiveness for nonlinear scattered data interpolation with multi resolution capabilities in high dimensions [13], [14], [15]. In our proposed approach, we harness the capabilities of RBFs to build a surrogate model of the energy consumption.

In our work, radial basis functions are used to estimation F in (1), as a linear combination of radial basis functions, as

$$F(X) \approx \sum_{i=1}^{M} f_i \phi(\|X - Y_i\|),$$

where the Y_i's are M center points used to define the M basis functions based on a kernel ϕ that defines the nature of the Radial Function used. The Y_i are chosen from a quasi-random (low-discrepancy) Sobol sequences [16] whose bounds are given by the minimal and maximal values of the temperature and consumption time-series. By doing so and unlike most RBF approaches in the literature, we avoid introducing hyper-parameters that lead to solving a nonlinear regression problem that may be ill-conditioned and requires solving a nonconvex optimisation problem.

Several experiments conducted with our data set showed that the Gaussian RBF kernel $\phi(r) = e^{-\varepsilon^2 r^2}$, with $r = \|X - Y\|$, provides very good results. ε denotes a shape parameter to be properly tuned to improve the problem conditioning. However depending on the application, other RBF can be used, such as the multiquadric RBF $\phi(r) = (1+\varepsilon^2 r^2)^{1/2}$, the inverse multiquadric RBF $\phi(r) = \dfrac{1}{(1+\varepsilon^2 r^2)^{1/2}}$, or the Poly Harmonic Splines $\phi(r) = r^{2k+1}$, with $k = 0, 1, 2....$

If we denote $X(t) = (C(t-1), C(t-2), ..., C(t-N_C), T(t), T(t-1), T(t-2), ..., T(t-N_T))$, as the vector of the combination of the consumption and temperature variables over a historical window defined by N_C and N_T, and θ, the set of the f_i's, the learning process consists in solving the following problem:

$$\min_{\theta} \sum_{t=1}^{N} (C(t) - \sum_{i=1}^{M} f_i \phi(\|X(t) - Y_i\|))^2. \quad (2)$$

This problem can be reformulated as the following linear regression problem

$$\min_{\theta} \|\bar{Y} - \bar{\Phi}\theta\|^2, \quad (3)$$

where $\bar{Y}$ is the vector of the N $C(t)$'s and $\bar{\Phi}$ is a (N,M) matrix in which each row is defined by $(\phi(\|X(t) - Y_1\|), \phi(\|X(t) - Y_2\|, ..., \phi(\|X(t) - Y_M\|))$.

The unique solution of problem (3) is given by

$$\theta = (\bar{\Phi}^T \bar{\Phi})^{-1} \bar{\Phi}^T \bar{Y}, \quad (4)$$

provided that $\bar{\Phi}$ has full rank.

An interesting feature of using predefined center points for RBFs is the possibility of training the model using a recursive Kalman Filter instead of using the batch approach (3), with solution (4). The potential advantage of using a Kalman Filter is its ability to take non stationary phenomena into account.

The estimation model is given by

$$\theta(t+1) = \theta(t) + v(t),$$
$$y(t) = \Phi(X(t))\theta(t) + w(t),$$

where $\Phi(X(t)) = (\phi(\|X(t) - Y_1\|), \phi(\|X(t) - Y_2\|, ..., \phi(\|X(t) - Y_M\|))$, $v(t)$ is the state noise, which is a Gaussian process with zero-mean and a covariance matrix Q, $w(t)$ is the output noise, which is also a Gaussian process with zero-mean and variance r.

The associated Kalman filter is given by

$$\hat{\theta}(t+1) = \hat{\theta}(t) + L(t)(y(t) - \Phi(X(t))\hat{\theta}(t)),$$
$$L(t) = P(t)\Phi(X(t))^T / (r + \Phi(X(t))P(t)\Phi(X(t))^T),$$
$$P(t+1) = Q + [I_d - L(t)\Phi(X(t))]P(t),\ P(0) = P_0,$$

where $P(t)$ is the covariance matrix of the estimation error $\theta(t) - \hat{\theta}(t)$ and P_0 , the covariance matrix of the initial error. In practice, P_0, Q can be chosen as diagonal matrices: $P_0 = bI_d$, $Q = qI_d$, with $b,q,r > 0$.

However, in this paper, only the batch approach (with the pseudo-inverse solution (4)) has been considered. The implementation of the Kalman Filter approach will be investigated in the future.

B. Feed Forward Neural Network Approach

In this paper, we will use 'Simple Artificial Neural Network - sANN' to refer to a feedforward neural network [17]. The inputs to this network consist of daily electricity consumption and temperature, while the cpredicted consumption is the single output of the network.

If we again denote $X(t) = (C(t-1), C(t-2), ..., C(t-N_C), T(t), T(t-1), T(t-2), ..., T(t-N_T))$, as the vector of the combination of the consumption and temperature variables over a historical window defined by N_C and N_T, the model of a feed forward neural network with $L-1$ hidden layers is given as follows:

$$z^0 = X(t)$$
$$z^k = \sigma(W^k z^{k-1} + b^k),\ 1 \le k \le L-1$$
$$C(t) = z^L = W^L z^{L-1} + b^L$$

Each output vector z^k of a layer k is defined in $\mathbb{R}^{n_k}$, where n_k denotes the number of neurons in layer k. $W^k \in \mathbb{R}^{n_k x n_{k-1}}$ denotes the weight matrix of neurons in layer k. $b^k \in \mathbb{R}^{n_k}$ is the offset vector of the k layer. σ is the activation function of the neurons (for example, *tanh*). The output vector z^L of the last layer L is the predicted consumption $C(t)$.

In this application, it appeared that using several hidden layers rather than a single hidden layer does not improve the quality of the prediction. Considering the computation time all the models with more than one hidden layer require more time to be performed. While the accuracy performance of the several layers model over training part is the same with the single layer models.

C. Recurrent Neural Network Approach

Recurrent Neural Networks (RNN) [18] represent a distinct group of neural networks characterized by their ability to process sequential data efficiently by retaining memory of past information through a feedback loop within the network. This memory allows RNNs to learn patterns based on context and make predictions on time series data. However, simple RNNs can struggle with long-term dependencies due to the vanishing gradient problem. [19] Recurrent Neural Networks with GRU layers are different form the SRNN model for the way how they pursue the information in the cell state. Their use come as a solution of the problem with Long-Short-Term-Memory layers in the state of the improving the weights in order to update the information for the produced result and the real values.

At time step t, information represented by X(t) enters the cell. Upon application of the sigmoid function in the reset gate, the information undergoes a bifurcation: a portion is transmitted through the forget gate, while the remaining portion, deemed significant, proceeds to the subsequent step. This significant information traverses through the cell state, contributing to the update of the cell by undergoing multiplication with the information derived from the previous cell state. Subsequently, the *tanh* function is applied to update the information based on new weights, which is then reintegrated into the cell state. The updated cell state is further modified through addition with the new information derived from the current step, thereby completing the information processing cycle within the GRU model as it is illustrated in schema 1.

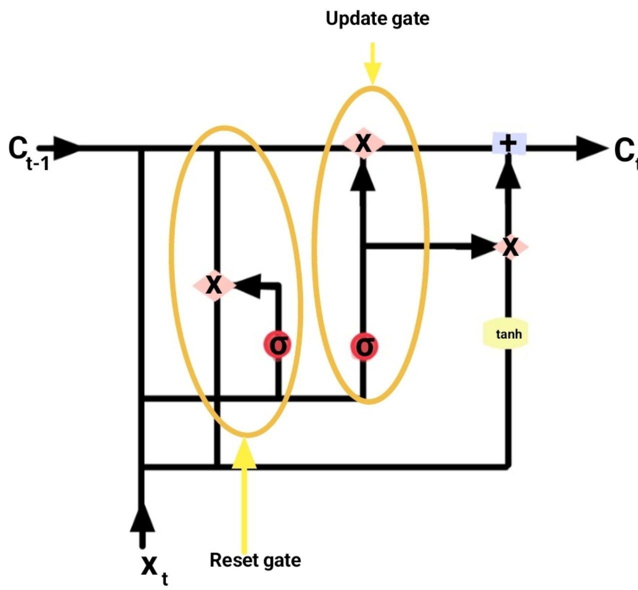

Fig. 1. RNN-GRU Schema

III. DATASET

Our study utilizes a real dataset sourced from Tirana, Albania, encompassing daily records spanning 34 months from March 2020 to December 2022. The dataset comprises two daily variables, including electricity consumption (MW/h) and average temperature (°C) . The figure below illustrates the relationship between two variables. The relation between temperature and consumption is positive during the summer period, indicating that higher temperatures correspond to increased consumption. While, during the winter periods of decreasing temperatures values of consumption are rising, representing a negative relation between them. The methodology applied to the dataset commenced with

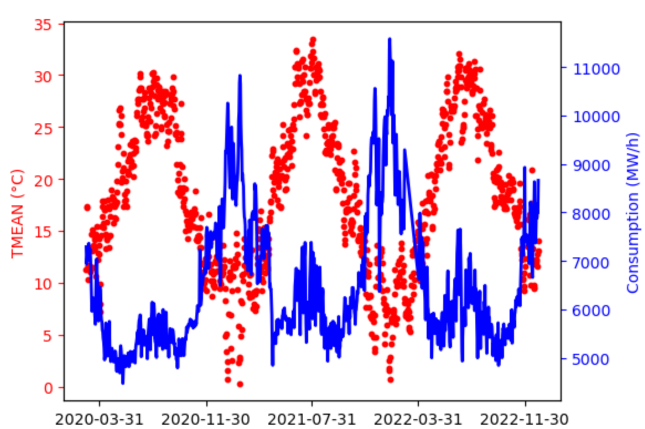

Fig. 2. Daily electricity consumption and average temperature relation

analyses employing techniques for handling missing data. We identified three days of missing temperature variables (11-02-2021, 12-02-2021, 20-02-2021) and one month (March 2022) of missing energy consumption data within the dataset. In both cases, linear interpolation models were employed to address missing data. These models were utilized to estimate the missing values based on the trend observed in the surrounding data points.

IV. RESULTS

In this section, we delve into the outcomes derived from our comprehensive analysis, which encompasses four distinct scenarios encompassing all scenarios examined. We evaluate the performance, computation time, and error measurements across these scenarios, providing a holistic understanding of their implications. First, we compare how the data used for training differs from the data used for testing. Second, we scrutinize the effect of varying historical windows for each variable, discerning their implications on predictive accuracy and robustness. By studying these scenarios, we aim to better understand how different factors affect the performance of our models. The primary objective of including two scenarios in training and testing datasets is to facilitate forecasting with a long-term prediction horizon, covering 10% of the data equal to 103 days, alongside short-term predictions made on a monthly basis.

TABLE I
SCENARIO COMBINATIONS

Scenario	N_C	N_T	Training	Testing
1	7	7	90% of data	10% of data
2	7	1	90% of data	10% of data
3	7	7	33 months	1 month
4	7	1	33 months	1 month

The models are applied over a scaled dataset using the standard scalar method.

For the neural network models, the hyperparameters are set as follows: 100 epochs, 150 neurons in the hidden layer, a learning rate of 0.005, and Tanh activation function. The batch size is set to 32, and the optimizer used is the Adam optimizer. These hyperparameter values are optimized through an error measurement between real data and predictions, considering intervals of [50, 100, 150] for epochs and number of nodes, and values of learning rate ranging from [0.002, 0.003, 0.005]. The activation function is selected from the performances of ReLU, SoftMax, and Sigmoid, and the loss is evaluated using mean squared error (MSE) technique.

For the RBF model, the shape parameter ε of the Gaussian kernel is set to 0.001, tuned within the interval of [0.0001 to 10], and the number of basis sets (M) is set to 16, tuned between 8, 16, 32, 64, 128, and 256. The construction of the RBF model involves utilizing a simple Sobol sequence for selecting center points, contributing to the optimization and effectiveness of the model.

The optimization process involves exploring the combinations of the hyperparameter space by assigning their values from predefined sets. The optimal set of hyperparameters is determined by selecting the combination of them that yields the lowest error between the predicted and real values.

The results from scenario 1, presented in Fig. 3 demonstrate how well the first model combination performed when tested on the dataset. Across all methods utilized, there is minimal deviation from the actual consumption values, indicating consistent performance.

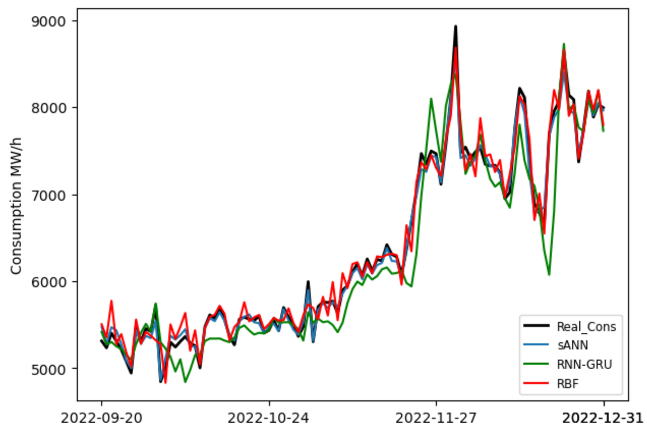

Fig. 3. Predicting over the 10% of the dataset, Scenario 1

In the scenario 2, Fig. 4, we notice that altering the historical data for temperature, focusing solely on the day prior to prediction, results in improved performance for the RNN model. However, this modification doesn't notably impact the performance of the sANN and RBF methods.

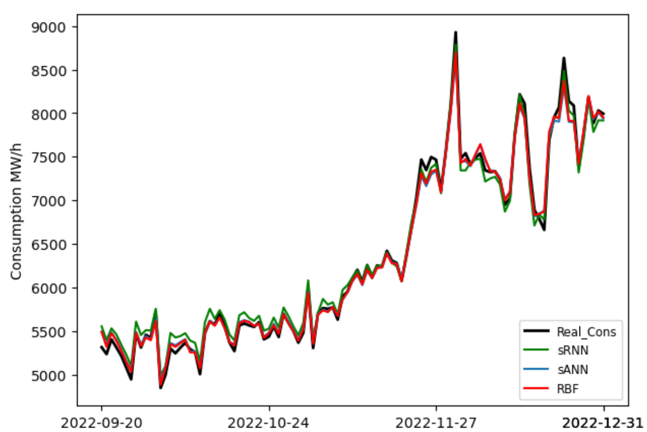

Fig. 4. Predicting over the 10% of the dataset, Scenario 2

The outcomes from scenario 3, presented in Fig.5, where the models predict over the last month of the dataset. Similar to scenario 1, in Fig.3, the sRNN model experiences comparable effects when incorporating the temperature data from the last 7 days.

From the last scenario presented in Fig. 6, it is evident that the performance of the RNN model has improved compared

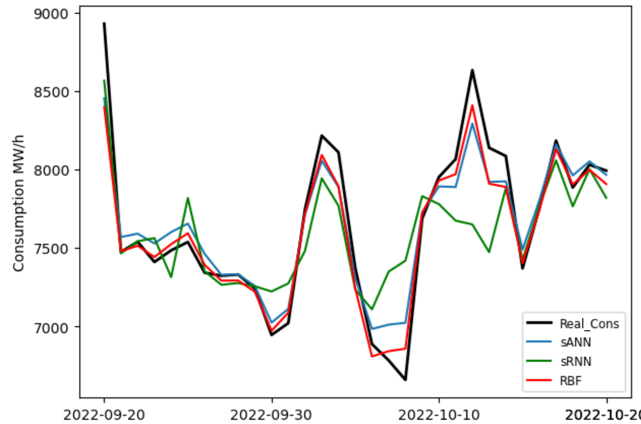

Fig. 5. Predicted results over the last month, Scenario 3

to same results from scenario 3, despite considering only the temperature value from the day preceding the prediction.

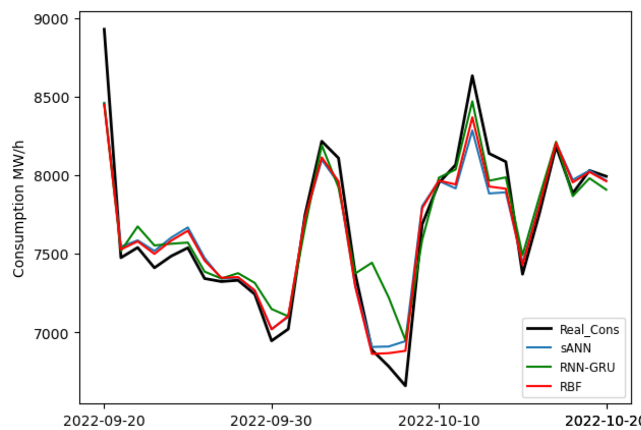

Fig. 6. Predicted results over the last month, Scenario 4

Assessing model performance in relation to computational time

Table II presents the computational performance for each method across all scenarios, measured in seconds. The first two scenarios consistently exhibit the same ordering of methods across all scenarios. However, the third and fourth scenarios reveal distinct performances among standard neural networks. Overall, the RBF method emerges as the top-performing model across all scenarios and methods.

TABLE II
MODEL COMPUTATION TIME (SEC)

Method	Scenario 1	Scenario 2	Scenario 3	Scenario 4
RNN-GRU	46.26	47.02	16.64	12.37
sRNN	25.43	22.45	53.95	16.96
sANN	14.26	11.06	27.42	21.69
RBF	0.78	0.84	1.54	1.44

In the last evaluation across all applied models, we analyze the error measurements computed against the real consumption data and the corresponding predicted results over this segment of the dataset.

The performance of these models was evaluated using the following error measurement metrics, while the variable c is the real electricity consumption and the $\hat{C}$ is the predicted electricity value:

- MAE (Mean Absolute Error):

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|C_i - \hat{C}_i|$$

- RMSE (Root Mean Squared Error):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(C_i - \hat{C}_i)^2}$$

- MAPE (Mean Absolute Percentage Error):

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{C_i - \hat{C}_i}{C_i}\right|$$

In the first scenario, we observe that RBF exhibits the lowest values of RMSE and MAPE. However, it is noteworthy that sANN achieves the lowest MAE value in this scenario.

For the second, third, and fourth scenarios, RBF consistently outperforms other models, demonstrating significantly lower values across all error measurements compared to alternative models.

TABLE III
ERROR METRICS

Model	Error Method	**sANN**	**sRNN**	**RNN-GRU**	**RBF**
Scenario 1	RMSE	1,475.0	1,443.6	1,500.8	158.0
	MAE	60.7	241.9	223.0	122.7
	MAPE	18.29	17.54	18.34	1.97
Scenario 2	RMSE	1,464.0	1,444.1	1,470.3	78.1
	MAE	63.1	84.7	921.9	55.0
	MAPE	18.18	18.01	16.09	0.83
Scenario 3	RMSE	649.9	610.8	624.5	138.2
	MAE	121.4	234.7	324.9	90.1
	MAPE	6.86	6.37	6.51	1.14
Scenario 4	RMSE	663.0	642.7	645.6	135.6
	MAE	110.4	300.8	127.5	97.3
	MAPE	6.99	6.82	6.80	1.24

V. CONCLUSIONS

This paper explores three machine learning approaches for forecasting electricity consumption, leveraging historical data while integrating temperature effects from the past 7 or 1 day(s). The scaling of historical data and testing models introduces a novel technique for adjusting variable importance within the model. Prioritizing fewer variables with significant impact results in three notable enhancements: simplification of the model, reduced reliance on extensive historical data, increased accuracy in results, and decreased computational time in model processing. Based on our findings, we recommend the adoption of the RBF model for electricity consumption prediction. The RBF model demonstrates superior performance in model fitting over real data, boasts lower computational time, and exhibits significant improvements in error measurement values. Additionally, the technical and theoretical implementation of the RBF approach is simpler compared to neural networks, as it relies on a linear regression problem rather than the inherently nonlinear regression problem associated with neural networks. Looking ahead, future work will explore the performance differences of these models in regions with diverse correlations between variables, and consider the incorporation of additional exogenous variables and new models. Additionally, other methods such as cross-validation for testing and training, as well as variations in prediction horizons, will be evaluated for their potential incorporation into these models.

ACKNOWLEDGMENT

Authors would like to express their gratitude to Institute of Geosciences and Distribution System Operator, Tirana, Albania, for providing the data.

REFERENCES

[1] "Electricity Regulatory Entity (ERE) - Albania." `https://www.ere.gov.al/en/`. Last Access date:29/03/2024.

[2] S. Ryu, J. Noh, and H. Kim, "Deep neural network based demand side short term load forecasting," *Energies*, vol. 10, no. 1, p. 3, 2016.

[3] F. Karanfil and Y. Li, "Electricity consumption and economic growth: Exploring panel-specific differences," *Energy policy*, vol. 82, pp. 264–277, 2015.

[4] H. Toros and D. Aydın, "Prediction of short-term electricity consumption by artificial neural networks using temperature variables," *European Journal of Science and Technology, EJOSAT*, no. 14, pp. 393–398, 2018.

[5] F. Apadula, A. Bassini, A. Elli, and S. Scapin, "Relationships between meteorological variables and monthly electricity demand," *Applied Energy*, vol. 98, pp. 346–356, 2012.

[6] T. R. Jena, S. S. Barik, and S. K. Nayak, "Electricity consumption & prediction using machine learning models," *Acta Tech. Corviniensis-Bull. Eng*, vol. 9, pp. 2804–2818, 2020.

[7] J. T. Connor, R. D. Martin, and L. E. Atlas, "Recurrent neural networks and robust time series prediction," *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 240–254, 1994.

[8] F. M. Bianchi, E. Maiorino, M. C. Kampffmeyer, A. Rizzi, and R. Jenssen, "An overview and comparative analysis of recurrent neural networks for short term load forecasting," *arXiv preprint arXiv:1705.04378*, 2017.

[9] M. Abumohsen, A. Y. Owda, and M. Owda, "Electrical load forecasting using lstm, gru, and rnn algorithms," *Energies*, vol. 16, no. 5, p. 2283, 2023.

[10] L. Babich, D. Svalov, A. Smirnov, and M. Babich, "Industrial power consumption forecasting methods comparison," pp. 307–309, 2019.

[11] J. Q. Wang, Y. Du, and J. Wang, "Lstm based long-term energy consumption prediction with periodicity," *energy*, vol. 197, p. 117197, 2020.

[12] E. Gjika and L. Basha, "A comparative study of statistical and deep learning models for energy load prediction," *International Journal of Electrical Engineering and Computer Science*, 2022.

[13] D. Broomhead and D. Lowe, "Multivariable functional interpolation and adaptive networks," *Complex Systems*, vol. 2, no. 3, pp. 321–355, 1988.

[14] P. Green and B. Silverman, "Nonparametric regression and generalized linear models," *Monographs on Statistics and Applied Probability*, 1994.

[15] A. Iske, "Multiresolution methods in scattered data modelling," *Lecture Notes in Computational Science and Engineering*, vol. 37, 2004.

[16] W. J. Morokoff and R. E. Caflisch, "Quasi-random sequences and their discrepancies," *Journal of Scientifc Computing*, vol. 16, no. 6, 1994.

[17] B. Yegnanarayana, *Artificial neural networks*. PHI Learning Pvt. Ltd., 2009.

[18] R. M. Schmidt, "Recurrent neural networks (rnns): A gentle introduction and overview," *arXiv preprint arXiv:1912.05911*, 2019.

[19] S.-H. Noh, "Analysis of gradient vanishing of rnns and performance comparison," *Information*, vol. 12, no. 11, p. 442, 2021.