Performance Analysis of Single and Multi-step Short-term Load Forecasts Using Multi-layer Perceptron

Athanasios Ioannis Arvanitidis
Electrical & Computer Engineering
University of Texas at San Antonio
San Antonio, USA
athanasios.arvanitidis@my.utsa.edu

Dimitrios Kontogiannis
Electrical & Computer Engineering
University of Thessaly
Volos, Greece
dimkonto@uth.gr

Georgios Vontzos
Electrical & Computer Engineering
University of Thessaly
Volos, Greece
gvontzos@uth.gr

Vasileios Laitsos
Electrical & Computer Engineering
University of Thessaly
Volos, Greece
vlaitsos@uth.gr

Dimitrios Bargiotas
Electrical & Computer Engineering
University of Thessaly
Volos, Greece
bargiotas@uth.gr

Miltiadis Alamaniotis
Electrical & Computer Engineering
University of Texas at San Antonio
San Antonio, USA
miltos.alamaniotis@utsa.edu

***Abstract*—Load forecasting is one of the most critical factors in modern power systems since it is the cornerstone for efficient monitoring, resource management and decision making. Therefore, there is an accrescent need for accurate and fast electrical load predictions. Many scientific approaches have been carried out in the field of load forecasting. In particular, the field of Machine Learning has attracted great research interest, due to the ability to adapt to time-series through forecasting tasks on multiple prediction horizons, a research area that presents challenges to several traditional methods. For this purpose, this research offers a thorough comparative study of several structural morphologies of Multi-Layer Perceptrons, in order to investigate electrical load forecasting accuracy for one, twelve and twenty-four time-steps ahead. Based on data from the Greek Power System for the years 2017 to 2019, the three proposed neural networks' structural morphologies are assessed in terms of precision through the Mean Absolute Error, Mean Squared Error, and Mean Absolute Percent Error of the predicted outcomes.**

***Index Terms*—multi-layer perceptron, univariate prediction, multivariate prediction, short-term load forecasting**

I. INTRODUCTION

Nowadays accurate load forecasting is crucial for power companies in order to sufficiently and reliably generate, transmit and distribute electric power. Non accurate load prediction could lead to limited facilities capacity, power supply shortage or even power interruption, and causes annoyance to stakeholders and consumers. Additionally, an accurate forecasting of electrical load supports proper infrastructures' maintenance and reduction of power companies' operational costs [1]. Load forecasting is commonly categorized as very short-term (VSLTF), short-term (SLTF), mid-term (MLTF) and long-term (LTLF) forecasting [2]. In this work, STLF is used to predict future demand.

In terms of electrical load forecasting several methodologies have been introduced and can be divided into two main categories the traditional and the modern methods. In traditional techniques, statistical methods are mainly utilized. These include models like Autoregression (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA), ARMA and ARIMA with exogenous inputs (ARMAX and ARIMAX respectively), Grey (GM) and Exponential Smoothing (ES) [3], [4].

On the other hand, modern load forecasting methods are considered machine learning, artificial intelligence-based and hybrid technics. Machine learning approaches include Support Vector Machine (SVM) models, which seems to be used extensively in forecasting issues. Additionally Artificial Neural Network (ANN) algorithms are very popular in recent years in time series prediction. The commonly used ANN algorithms for electrical load forecasting are Recurrent Neural Networks (RNN), such as Long-Short Term Memory (LSTM), Convolutional Neural Networks (CNN) and Feed-Forward Neural Networks, like Multi-Layer Perceptron (MLP) [5], [6].

Several recent studies propose the use of MLP models in future electrical demand prediction [7]. Although MLPs are the simplest type of ANN, they are used to complex non-linear problems. They not only perform well with large number of input data but also provide fast predictions after training [8], such as Kontogiannis et al. were observed at experiments with residential real world data compared with LSTM and CNN models [9]. Arvanitidis et al. used MLP models to propose novel train data pre-processing strategies [10] and clustering techniques [11] for SLTF. Furthermore, MLP architecture is extended to conduct a day-ahead electricity price forecasting [12]. MLP is impactful in short term

 ISBN: 978-1-68558-054-4

forecasting tasks involving not only load related time series but also wind power forecasting. In [13], the Simulated Annealing optimization algorithm is used to specify the hyperparameters of the under investigation regressor models. The MLP model seems to predict wind power accurately compared to the other forecasting methods examined.

Electric load forecasting is a vital step in the planning of the power system industry. It is crucial to the management of the power system and the scheduling of electricity in order to ensure the system's cost-effective and uninterrupted performance. As a result, it offers multiple significant benefits for controlling generating capacity, scheduling, management, peak reduction, market assessment, and demand response. As a result, forecasting on diverse time horizons has shown to be incredibly effective in meeting the various criteria of their application. The literature provides a plethora of publications that investigate the subject of load forecasting in various prediction horizons, but without indicating which is the optimum and most efficient. Hence, the extensive comparison of univariate and multivariate short-term load forecasting methods is the novelty of this work. In this paper, three different optimized structural morphologies of MLPs are applied to anticipate load values for one, twelve, and twenty-four hours ahead, respectively. In order to determine which of the suggested morphologies of the neural networks delivers more accuracy, the results are compared based on the various prediction metrics.

The paper is organized as follows. Section II analyzes the necessary actions related to the preprocessing of the data used for forecasting, while Section III analyzes the forecasting model. Section IV describes the hyperparameters' optimisation algorithm, proposes three different MLP prediction models and presents the simulation results of short-term load forecasting, while Section V concludes the results of the paper.

II. Dataset Overview

The datasets utilized in this research project for the performance evaluation of the multilayer perceptron structure on total electricity load predictions for the Greek power system consist of hourly measurements of total load in MW, temperature in °C and relative humidity expressed as a percentage. The samples in this dataset cover a three year time period spanning from 2017-01-01 to 2019-12-31. The samples of total load were made publicly available by the transparency platform Entso-E [14] and the environmental variables of temperature and relative humidity were accessed through the MERRA-2 research and analysis platform [15].

Since this project considers the short-term forecasting tasks of 1, 12 and 24 hour-ahead total load prediction, the features utilized in this analysis include the temporal variables for the hour and day of the week encoded in the value intervals 0-23 and 1-7 respectively, the temperature and humidity variables for the target time interval and the historical load features corresponding to the same time intervals for the previous 7 days. The resulting datasets did not have any missing values and the values were scaled through min-max normalization [16]. The dataset was split into a training set containing 80% of samples and a test set containing 20% of samples based on common practices with regards to the splitting ratio [17]. Figure 1 presents the target variable of load.

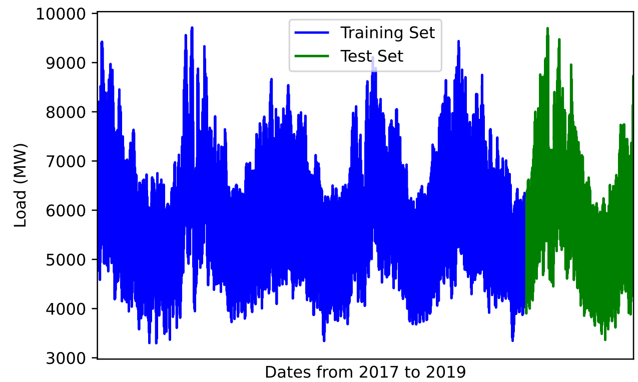

Fig. 1. Graphical representation of the dataset load values.

III. Multi-Layer Perceptron

The MLP is a neural network structure that belongs to the class of feed forward artificial neural network architectures as it expresses a fully connected acyclic computation graph that focuses on the task of function approximation in order to derive a model that efficiently predicts the target variables given the input features x through the target function that describes the relationship between the inputs and the output, defined as f. This neural network consists of neurons that perform the computations following the perceptron supervised learning algorithm [18]. These neurons are organized in layers that express different roles in the computation path. The input layer receives the features from the initial input dataset, the hidden layers express the transition from the input to the output through a series of computations involving the adjustment of a weight matrix W that quantifies the importance of the input features towards the prediction of the target output and a bias vector b that is used to offset the computed results of the neurons. The output of each hidden layer is determined based on activation functions that evaluate the importance of computations and select the data that will proceed deeper in the network. The output layer derives the estimated values from the output of the last hidden layer. Consequently, the estimated values $f(x)$ derived from a multi-layer perceptron with one hidden layer could be expressed by formula (1) denoting the subsequent adjustments to weight matrix W^1 and bias vector b^1 for the transition from the input to the first hidden layer through the activation function as well as the application of the activation function and the adjustments to the weight matrix W^2 and bias vector b^2 for the transition from the hidden layer to the output layer.

$$f(x) = G(b^2 + W^2(s(b^1 + W^1 x))) \tag{1}$$

Since this study examines univariate and multivariate load forecasting tasks, it is worth noting that the MLP is modified accordingly in order to derive the appropriate estimated values. Therefore, the number of neurons in the output layer needs to match the number of predicted output variables, corresponding to the target time series. The neural networks with a single output neuron are known as univariate MLPs and are used for hourly load forecasting, while the MLPs with several neurons in the output layer are known as multivariate and are employed in day ahead load prediction, as Figure 2 depicts. The MLP structure is typically trained through back propagation with gradient descent. Impactful structural parameters that could affect the training process and the performance of the model include the number of neurons at each layer, the number of hidden layers and the number of training epochs. Additionally, learning parameters, such as learning rate, the types of activation functions and the optimizer could be equally important to the generalization capabilities of the model [19].

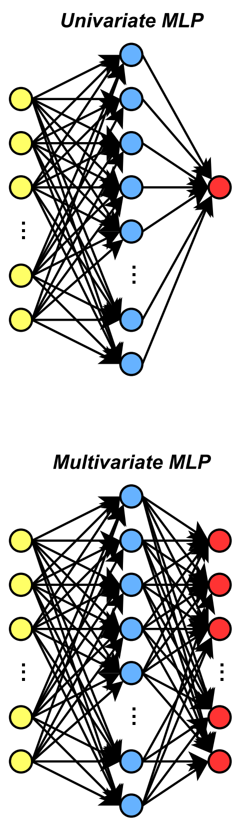

Fig. 2. Structural difference between univariate and multiariate MLPs.

IV. Evaluation Metrics

In this section, the metrics used for the performance evaluation of the single and multi-step forecasting models are outlined in order to further explain their role in our experiments. Firstly, Mean Absolute Error (MAE) is utilized as a simple and interpretable metric in order to naturally describe the average error of the MLP. Secondly, Mean Squared Error (MSE) is included in the examination of model performance since it is a scale dependent error metric that considers the direction of the predicted values. Lastly, Mean Absolute Percentage Error (MAPE) was utilized in order to denote the generalized relative error of the models. Given the forecasted data points y_i and the actual values x_i in a set of n observations, MAE, MSE and MAPE are given by (2), (3) and (4) respectively [20].

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \tag{2}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2 \tag{3}$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}|\frac{y_i - x_i}{x_i}| \tag{4}$$

V. Results

This study investigates the implications of MLPs' univariate and multivariate prediction procedures on the precision of short-term load forecasting. Also, the variations in the execution time of a brute force optimization algorithm, which is employed for the optimal hyperparameter selection on the various morphologies of the proposed MLPs, are thoroughly addressed. The suggested MLPs were developed using Python's scikit-learn framework [21], while the computer system utilized has an Intel Core i7-4510U CPU running at 2.00 GHz and 8 GB of installed memory.

Initially, the suggested brute force optimization approach is used to determine the ideal values for two major neural network hyperparameters: the number of neurons in the hidden layer and the number of epochs throughout the training process. For the sake of simplicity, each suggested MLP comprises of a hidden layer, as it is adequate for the load forecasting issue. Most neural networks used for load forecasting include only one hidden layer in order to reduce computational complexity while yet providing fast online results.

In this paper, we examine the structural morphology of three neural networks in order to estimate hourly load values for one hour, twelve hours, and twenty-four hours ahead, respectively. Thus, in pursuance of a direct comparison on the hyperparameter selection, we suggest that the number of neurons in the hidden layer is determined as a function of the number of neurons in the input layer for each possible structural morphology. In each case examined, the minimum acceptable number of hidden neurons is half of the number of neurons in the input layer, while the maximum number of hidden neurons can reach up to three times the number of input neurons. The ideal number of epochs, on the other hand, is derived by sequential scanning within the closed interval [200, 2000], where 200 is the least permitted limit of iterations and 2000 is the maximum limit of epochs.

In the following subsections, the calculation of the hyperparameters of the three different proposed MLPs, as well as the execution time of the brute force optimization approach for each case and their performance evaluation, in terms of accuracy, for short-term load forecasting are thoroughly examined.

A. One hour ahead load forecasting - Univariate Neural Network

The case of a univariate MLP, i.e., a neural network with a single neuron in the output layer, is first investigated. This network generates the hourly value of the load for the following hour. The input data to this MLP, which consists of 11 neurons in the input layer, is as follows:

- A label for the time for which the forecast is being performed, represented as an integer within [0, 23].
- An integer that belongs to the range [1, 7], serving as a label to identify the day being predicted. Sunday is represented by the value 1, Monday by the value 2, etc.
- The hourly temperature value for the precise time of day for which the prediction is conducted.
- The hourly humidity estimation for that particular hour of day in which the prediction is made.
- Seven hourly load values for the period from the current time up to one week in beforehand of the prediction.

B. Twelve hours ahead load forecasting - Multivariate Neural Network

The case of an MLP used for 12 hours ahead load forecasting is then considered, i.e. a neural network with 12 neurons in the output layer. In this scenario, the number of neurons in the input layer is 109 and results from the following input data:

- An integer within the range [1, 7], serving as a label to identify the day being predicted. Sunday is represented by the value 1, Monday by the value 2, etc.
- A vector consisting of 12-hourly temperature values for specific hours of the day for which the prediction is conducted. A day is divided into two instances in the dataset. Thus, the first vector of the day concerns the hours from midnight to 11 am and contains the corresponding hourly temperature data.
- A vector consisted of 12-hourly humidity estimations, similar to the case of temperature, for that particular hour of day in which the prediction is made.
- A vector of 84-hourly load values concerning the period from the current time up to one week in beforehand of the prediction, respectively.

C. Twenty-four hours ahead load forecasting - Multivariate Neural Network

The last case study focuses at a multivariate MLP that is used to estimate day-ahead load and has 24 neurons in the output layer, one for each hour of the day. Similar to the other examples, the number of input data determines the quantity of neurons in the input layer. Thus, in this case, the 217 input neurons result from the following data:

- An integer in the range [1, 7], acting, as in the earlier cases, as a label to designate the day being forecast.
- A vector consisting of 24-hourly temperature values for the day of which the prediction is conducted.
- A vector consisted of 24-hourly humidity values for the day of which the prediction is conducted.
- A vector of 168-hourly load values concerning the period from the current time up to one week in beforehand of the prediction, respectively.

The boundaries of the hyperparameters for each MLP utilized in each case study are reported in Table I. The results of the optimization process used to identify the ideal hyperparameters for the MLPs of each case study are summarized in Table II. Subsequently, the optimized neural networks are used for load prediction and their results are compared, in terms of accuracy, in order to decide whether the univariate or the multivariate structural morphology responds better to the STLF issue. Furthermore, Table III compiles the findings of the MAE, MSE, and MAPE metrics yielded from each case study. Lastly, Figure 3 and Figure 4 graphically represent the STLF outcomes for each prediction method considered.

TABLE I
RESULTS OF THE OPTIMIZATION APPROACH FOR EACH CASE STUDY.

	Boundaries		**Step**	
MLP	***Neurons***	***Iterations***	***Neurons***	***Iterations***
1h Ahead	[6, 33]	[200, 2000]	6	10
12h Ahead	[55, 327]	[200, 2000]	55	10
24h Ahead	[109, 651]	[200, 2000]	109	10

TABLE II
RESULTS OF THE OPTIMIZATION APPROACH FOR EACH CASE STUDY.

MLP	***Neurons***	***Iterations***	***Time (H:MM:SS)***
1h Ahead	33	2000	2:20:11
12h Ahead	275	1800	0:56:59
24h Ahead	436	1800	1:19:43

TABLE III
ACCURACY METRICS DERIVED FROM EACH MLP FOR STLF.

MLP	***MAE***	***MSE***	***MAPE (%)***
1h Ahead	182.076	67603.22	2.774
12h Ahead	162.845	54383.46	2.435
24h Ahead	187.315	66564.93	2.742

VI. CONCLUSION

In this paper, a detailed comparative analysis was conducted with the purpose of providing accurate load forecasting results for the Greek Power System for the period 2017-2019. The implementation of three different structural morphologies of MLP models were developed and assessed.

Based on the results, it is concluded that load forecasting with the implementation of MLPs plays a critical role in the safety, stability, and sustainability of modern energy systems. More specifically, it is observed that for all three cases, the results are quite satisfactory. The error values for the one hour ahead, and twenty four hours ahead forecast are very similar in terms of error metrics. The twelve hour ahead model exhibited improved performance compared to the other forecasting horizons. This might be because the number of

 ISBN: 978-1-68558-054-4

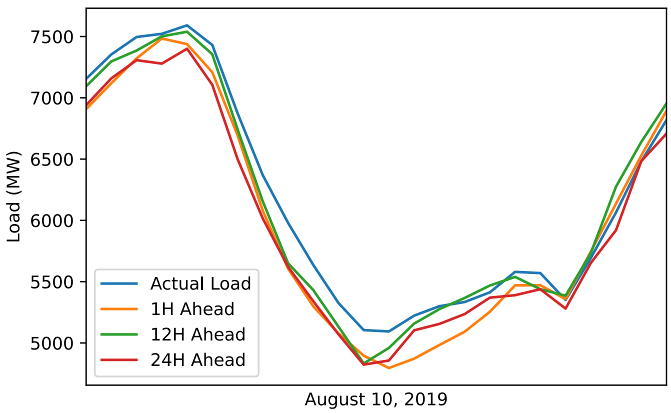

Fig. 3. Graphical comparison of the load prediction results of each considered method for a whole day in August.

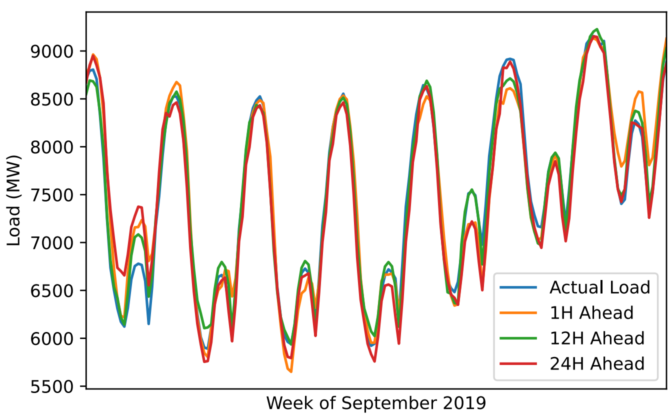

Fig. 4. Graphical comparison of the load forecasting results of each considered method, for a week in September.

input layer neurons, and hence the quantity of data utilized as input data, more closely approximates the case study under discussion. Given the data quality and data seasonality of the dataset, all models yielded relatively low error metrics. This observation proves the correct use of the algorithm in a time-series with multiple seasonalities, like the one studied. Since, the most accurate forecast was found for the case of twelve hours ahead, the ability of the algorithm to adapt to multi-step ahead forecasting is highlighted. Finally, it is worth noting that these models could assist the uninterrupted and reliable operation of Smart Grids using real-time data, where day ahead load forecasting delivers significant value.

In future work, this model could be evaluated on more complex load forecasting issues and compared with other deep learning models, used as benchmarks. Also, the proposed techniques could be applied to Demand Side Management and Demand Response programs [22], which have developed rapidly in recent years due to the global increase of energy consumption.

REFERENCES

[1] E. A. Feinberg and D. Genethliou, "Load forecasting," in *Applied mathematics for restructured electric power systems*. Springer, 2005, pp. 269–285.

[2] A. A. Mamun, M. Sohel, N. Mohammad, M. S. Haque Sunny, D. R. Dipta, and E. Hossain, "A comprehensive review of the load forecasting techniques using single and hybrid predictive models," *IEEE Access*, vol. 8, pp. 134 911–134 939, 2020.

[3] M. A. Hammad, B. Jereb, B. Rosi, and D. Dragan, "Methods and models for electric load forecasting: a comprehensive review," *Logistics, Supply Chain, Sustainability and Global Challenges*, vol. 11, no. 1, pp. 51–76, 2020.

[4] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *International Journal of Forecasting*, vol. 32, no. 3, pp. 914–938, 2016.

[5] I. K. Nti, M. Teimeh, O. Nyarko-Boateng, and A. F. Adekoya, "Electricity load forecasting: A systematic review," *Journal of Electrical Systems and Information Technology*, vol. 7, no. 1, pp. 1–19, 2020.

[6] A. Román-Portabales, M. López-Nores, and J. J. Pazos-Arias, "Systematic review of electricity demand forecast using ann-based machine learning algorithms," *Sensors*, vol. 21, no. 13, p. 4544, 2021.

[7] A. Gasparin, S. Lukovic, and C. Alippi, "Deep learning for time series forecasting: The electric load case," *CAAI Transactions on Intelligence Technology*, vol. 7, no. 1, pp. 1–25, 2022.

[8] "Deep Learning Tutorial for Beginners: Neural Network Basics." [Online]. Available: https://www.guru99.com/deep-learning-tutorial.html#5

[9] D. Kontogiannis, D. Bargiotas, and A. Daskalopulu, "Minutely active power forecasting models using neural networks," *Sustainability*, vol. 12, no. 8, p. 3177, 2020.

[10] A. I. Arvanitidis, D. Bargiotas, A. Daskalopulu, V. M. Laitsos, and L. H. Tsoukalas, "Enhanced short-term load forecasting using artificial neural networks," *Energies*, vol. 14, no. 22, p. 7788, 2021.

[11] A. I. Arvanitidis, D. Bargiotas, D. Kontogiannis, I. P. Panapakidis, and L. H. Tsoukalas, "Clustering informed mlp models for fast and accurate short-term load forecasting," *Energies*, vol. 15, no. 4, p. 1295, 2022.

[12] D. Kontogiannis, D. Bargiotas, A. Daskalopulu, A. I. Arvanitidis, and L. H. Tsoukalas, "Error compensation enhanced day-ahead electricity price forecasting," *Energies*, vol. 15, no. 4, p. 1466, 2022.

[13] A. I. Arvanitidis, D. Kontogiannis, G. Vontzos, V. Laitsos, and D. Bargiotas, "Stochastic heuristic optimization of machine learning estimators for short-term wind power forecasting," in *2022 57th International Universities Power Engineering Conference (UPEC)*. IEEE, 2022, pp. 1–6.

[14] "Central collection and publication of electricity generation, transportation and consumption data and information for the pan-european market." [Online]. Available: https://transparency.entsoe.eu/load-domain/r2/totalLoadR2/show?name=

[15] "Merra-2." [Online]. Available: https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/

[16] S. K. Patro and K. K. sahu, "Normalization: A preprocessing stage," *IARJSET*, p. 20–22, 2015.

[17] "Split training and testing data sets in python," Oct 2020. [Online]. Available: https://www.askpython.com/python/examples/split-data-training-and-testing-set

[18] "Dive into deep learning." [Online]. Available: http://d2l.ai/chapter multilayerperceptrons/mlp.html

[19] [Online]. Available: https://towardsdatascience.com/what-are-hyperparameters-and-how-to-tune-the-hyperparameters-in-a-deep-neural-network-d0604917584a

[20] "Error metrics: How to evaluate your forecasting models," May 2022. [Online]. Available: https://www.jedox.com/en/blog/error-metrics-how-to-evaluate-forecasts/

[21] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, D. Passos, A. and Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[22] V. M. Laitsos, D. Bargiotas, A. Daskalopulu, A. I. Arvanitidis, and L. H. Tsoukalas, "An incentive-based implementation of demand side management in power systems," *Energies*, vol. 14, no. 23, p. 7994, 2021.