Short-Term Electric Load Demand Forecasting on Greek Energy Market using Deep Learning: A comparative study.

1st George Emmanouilidis
Dept. of Informatics
Aristotle University of Thessaloniki
Thessaloniki, Greece
georemma@csd.auth.gr

2nd Maria Tzelepi
Dept. of Informatics
Aristotle University of Thessaloniki
Thessaloniki, Greece
mtzelepi@csd.auth.gr

3rd Anastasios Tefas
Dept. of Informatics
Aristotle University of Thessaloniki
Thessaloniki, Greece
tefas@csd.auth.gr

***Abstract*—In this paper, we deal with the short-term Electric Load Demand Forecasting problem, considering the Greek Energy Market. Particularly, we focus on two short-term cases, namely one-hour-ahead forecasting and one-day-ahead forecasting. The objective of this paper is to provide a comparative study considering the aforementioned problems on Greek Energy Market. To this end, we implement a wide variety of deep learning models (e.g., MLPs, CNNs, LSTMs, GRUs) that have been utilized in the recent literature either considering the ELDF task or generic time-series forecasting tasks. The extensive experimental evaluation has led to useful remarks.**

***Index Terms*—Energy load demand forecasting, Greek energy market, Short-term, One-hour-ahead, One-day-ahead, Deep Learning.**

I. Introduction

Electric Load Demand Forecasting (ELDF) refers to the challenging task of predicting the electricity demand by observing historical load data [1]–[3]. Predicting the load demand in an area is of utmost importance in power industry, since it is linked with many essential applications ranging from power system operation and planning to energy trading [4].

Earlier works for addressing the task of ELDF include statistical models [5], [6] and machine learning models [7], while later works focus on Deep Learning (DL) models [8], [9], following their successful application on various computer vision problems [10]–[13].

Different DL models have been proposed for tackling the ELDF problem in the literature, including either Multi-Layer Perceptron (MLP) architectures [14] or Recurrent Neural Networks (RNN) with Long Short-Term Memory (LSTM) modules [15].

Three categories of ELDF can be discriminated based on the time-scale, namely short-term forecasting which concerns forecasting of a few hours up to one-day ahead or a week ahead [16], mid-term load forecasting which concerns a forecasting of a week to one year ahead [17], and the long-term forecasting with a time frame of up to several years ahead [18].

In this paper, we deal with ELDF problem considering the Greek Energy Market, which has been studied to some extent during the recent years [14], [19]–[22]. More specifically, we investigate two particular short-term cases, i.e., one-hour-ahead load demand forecasting, which is a quite well studied problem [23]–[27], as well as the one-day-ahead forecasting problem [14], [20]. Our target is to provide a comparative study considering the aforementioned problems on Greek Energy Market, utilizing various DL models (e.g., MLPs, CNNs, LSTMs or the lighter GRUs) that have been utilized in the recent literature either considering the ELDF task or generic time-series forecasting tasks, e.g., [28].

The remainder of the manuscript is structured as follows. First, the investigated DL models are presented in Section II. Subsequently, in Section III the experiments conducted in order to evaluate the used models are provided. Finally, some conclusions are drawn in Section IV.

II. Deep Learning Models

In this paper we deal with the problem of short-term ELDF considering the Greek Energy Market. Particularly, we deal with one-hour-ahead and one-day-ahead forecasting. Our goal is to provide a comparative study of the most popular DL models used in the literature either considering the ELDF task, or generic time-series forecasting tasks.

To do so, we first tackle ELDF as a typical univariate time-series forecasting task naturally utilizing RNN-based models. More specifically, we first use simple CNN models with one or two layers, providing also experiments with various padding techniques. We also develop various hybrid CNN-LSTM and CNN-GRU models [28], as well as simple single-layered LSTM and GRU models. Furthermore, two-layered LSTM and GRU models are used. Additionally, we explore an attention mechanism [29]. Specifically, a self-attention mechanism is applied on the output of the LSTM and GRU models. In addition, a Bidirectional LSTM (BiLSTM) based a Seq2Seq Encoder-Decoder architecture is also investigated. Finally, an LSTM-based and a GRU-based Seq2Seq Encoder-Decoder architecture are also developed for addressing the ELDF task.

Apart from the typical time-series forecasting approach with RNN-based models, we also use lightweight MLP and CNN models with specific input features (i.e., past load demand along with temperature information) as it will be explained. In this case, we use MLP and CNN models. More specifically, we first use the MLP model proposed in [20], while considering the one-day-ahead forecasting task, another version is also developed using the same input features, and 24 MLP models each one tasked with predicting a different hour. Finally, single and two-layered CNNs are used.

III. Experimental Evaluation

In this Section, we first present the dataset used in the performed experiments, followed by the evaluation metric. Subsequently, the implementation details, followed by the experimental setup is provided. Finally, the experimental results are presented.

A. Dataset

In this work, we use past load data provided by the *Greek Public Power Corporation*. We also use weather information (i.e., temperature in Thessaloniki) derived from OpenWeather[1]. Particularly, we use 6 years of data for the model's training, that is load and temperature data for years 2012-2016, for validation load and temperature data for the year 2017, while for testing we use data for the year 2018.

B. Evaluation Metrics

Mean Absolute Percentage Error (MAPE) is used as evaluation metric. MAPE for a set of n test samples is defined as follows:

$$\mathrm{MAPE} = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{\mathbf{y}_t - \hat{\mathbf{y}}_t}{\mathbf{y}_t} \right|, \qquad (1)$$

where $\mathbf{y}_t$ is the ground truth and $\hat{\mathbf{y}}_\mathbf{t}$ is the model's prediction.

C. Implementation Details

All the models were trained for 2,000 epochs, using mini-batch gradient descent with a mini-batch of 128 samples. The learning rate is set to 0.003. The models are trained on an NVIDIA GeForce GTX 1050Ti with 4GB of GPU memory. Mean Square Error (MSE) loss was used for training. All the models were implemented using Tensorflow [30].

D. Experimental Setup

Two sets of experiments were conducted where in the former one we deal with ELDF as a typical univariate time-series forecasting task, using naturally RNN-based models. The last 10 days (considering the load consumption) are used to construct the time-steps of our input, in this case. In the latter one we use lightweight MLP and CNN models, and design the input features. More specifically, similarly to previous approaches [14], [20] we use as input 171 features comprising of the load of previous day, load of the day a week before and of the day a month before the day whose load demand we want to predict. Furthermore, temperature of the previous day, of the day a week before, and of the day a month before is used, as well as temperature of the day whose load demand we want to predict. Finally, two binary indicators for weekend and holiday are used, while an indicator of which day of the week is the day whose load demand we want to predict is also used. In both cases we consider the one-hour-ahead forecasting and the one-day-ahead forecasting tasks.

E. Experimental Results

First, in Table I the experimental results in terms of MAPE considering the typical time-series forecasting approach on the one-hour-ahead load demand forecasting task are provided, along with the details considering the layers and filters/neurons, as well as the variations on padding (causal padding is tested since it is a common choice considering time-series). As it is demonstrated, better performance is accomplished using two-layered LSTM and GRU. As expected the CNN implementations do not perform well in this scenario. This is attributed to the inability to remember past observations. Regarding the Seq2Seq architectures, better results are given by the BiLSTM-based endoder-decoder. Finally, we can observe that we can improve the results on LSTM and GRU by applying attention.

Subsequently, in Table II the corresponding results considering the one-day-ahead task are provided. As it is shown, better performance achieves the two-layered GRU. In general, MAPE, as expected is higher as compared to the one-hour-ahead problem, except for the CNN architectures. The Seq2Seq implementations perform poor on the one-day-ahead forecasting task, in opposition to the one-hour-ahead task, which needs further investigation. Causal padding, similarly to the previous case, generally do not provide improvements. In this case, we finally observe that attention do not provide remarkable changes.

Next, in Table III the experimental results in terms of MAPE considering the designed input features using simple and lightweight MLP and CNN models approach on the one-hour-ahead load demand forecasting task are provided, along with the details considering the layers and filters. It should be highlighted that these features are designed to tackle the one-day-ahead problem [20], however we also apply them in this considered scenario, too. As it is shown, better performance is accomplished by the 1-layered CNN without padding. The performance is generally poorer as compared to the typical time-series approach considering the one-hour-ahead problem, which is attributed to the fact that the input features are designed for the one-day-ahead task.

Finally, in Table IV the corresponding results considering the one-day-ahead problem are provided. As it is demonstrated, better performance is accomplished using a simple 2-layered CNN. Furthermore, it is shown that considering multiple MLPs we can achieve significantly performance as compared to the single MLP. However, this comes with significant increase in training time. Finally, we can notice that considering the one-day-ahead problem, we can achieve better

[1] https://openweathermap.org/

performance considering this approach, as compared to the typical time-series approach. However, this can be attributed to the utilization of the temperature information in the former case, as well as to the capture of a longer period of past load.

TABLE I
MODEL PERFORMANCE IN TERMS OF MAPE (%) CONSIDERING THE TYPICAL TIME-SERIES FORECASTING APPROACH ON THE ONE-HOUR-AHEAD LOAD DEMAND FORECASTING TASK.

Method	Padding	MAPE (%)
1-layered CNN(32) [31]	no padding	13.97
1-layered CNN(32) [31]	causal	16.52
2-layered CNN(32) [31]	no padding	15.12
2-layered CNN(32) [31]	causal	14.87
2-layered CNN(32) + LSTM(16) [28]	no padding	2.11
2-layered CNN(32) + LSTM(16) [28]	causal	2.20
2-layered CNN(32) + GRU(16) [28]	no padding	2.30
2-layered CNN(32) + GRU(16) [28]	causal	2.17
1-layered LSTM(32)	-	2.51
1-layered GRU(32)	-	4.21
1-layered GRU(64)	-	2.19
2-layered LSTM(32, 16) [32]	-	**1.99**
2-layered GRU(32, 16) [32]	-	**1.99**
1-layered LSTM(32) + Attention [29]	-	2.15
1-layered GRU(32) + Attention [29]	-	2.30
BiLSTM Seq2Seq [33]	-	2.04
1-layered LSTM Seq2Seq	-	2.34
1-layered GRU Seq2Seq	-	2.28

TABLE II
MODEL PERFORMANCE IN TERMS OF MAPE (%) CONSIDERING THE TYPICAL TIME-SERIES FORECASTING APPROACH ON THE ONE-DAY-AHEAD LOAD DEMAND FORECASTING TASK.

Method	Padding	MAPE (%)
1-layered CNN(32) [31]	no padding	6.94
1-layered CNN(32) [31]	causal	7.01
2-layered CNN(32) [31]	no padding	5.99
2-layered CNN(32) [31]	causal	7.17
2-layered CNN(32) + LSTM(16) [28]	no padding	5.63
2-layered CNN(32) + LSTM(16) [28]	causal	5.75
2-layered CNN(32) + GRU(16) [28]	no padding	5.74
2-layered CNN(32) + GRU(16) [28]	causal	5.85
1-layered LSTM(32)	-	5.83
1-layered GRU(32)	-	5.66
1-layered GRU(64)	-	5.72
2-layered LSTM(32, 16) [32]	-	5.64
2-layered GRU(32, 16) [32]	-	**5.52**
1-layered LSTM(32) + Attention [29]	-	5.81
1-layered GRU(32) + Attention [29]	-	5.63
BiLSTM Seq2Seq [33]	-	22.81
1-layered LSTM Seq2Seq	-	22.80
1-layered GRU Seq2Seq	-	22.81

TABLE III
MODEL PERFORMANCE IN TERMS OF MAPE (%) CONSIDERING THE DESIGNED INPUT ON THE ONE-HOUR-AHEAD LOAD DEMAND FORECASTING TASK.

Method	Padding	MAPE (%)
MLP [20]	-	21.57
1-layered CNN(32)	no padding	**5.72**
1-layered CNN(32)	causal	5.76
2-layered CNN(32)	no padding	7.69
2-layered CNN(32)	causal	6.51

TABLE IV
MODEL PERFORMANCE IN TERMS OF MAPE (%) CONSIDERING THE DESIGNED INPUT ON THE ONE-DAY-AHEAD LOAD DEMAND FORECASTING TASK.

Method	Padding	MAPE (%)
MLP [20]	-	5.80
Multiple MLPs	-	3.51
1-layered CNN(32)	no padding	3.79
1-layered CNN(32)	causal	3.62
2-layered CNN(32)	no padding	**3.29**
2-layered CNN(32)	causal	3.62

IV. CONCLUSIONS

In this paper we dealt with one-hour-ahead and one-day-ahead Electric Load Demand Forecasting problems, considering the Greek Energy Market. Our target is to provide a comparative study considering the aforementioned problems on Greek Energy Market. A wide variety of state-of-the-art deep learning models, considering time-series forecasting problems, were implemented, leading to useful remarks.

ACKNOWLEDGMENT

This work is co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH - CREATE - INNOVATE (project code: T2EDK-03048).

REFERENCES

[1] N. Ahmad, Y. Ghadi, M. Adnan, and M. Ali, "Load forecasting techniques for power system: Research challenges and survey," *IEEE Access*, vol. 10, pp. 71 054–71 090, 2022.

[2] 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.

[3] N. Passalis and A. Tefas, "Global adaptive input normalization for short-term electric load forecasting," in *Proceedings of the 2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2020, pp. 1–8.

[4] M. Jacob, C. Neves, and D. Vukadinović Greetham, *Forecasting and assessing risk of individual electricity peaks*. Springer Nature, 2020.

[5] S. S. Pappas, L. Ekonomou, P. Karampelas, D. Karamousantas, S. Katsikas, G. Chatzarakis, and P. Skafidas, "Electricity demand load forecasting of the hellenic power system using an arma model," *Electric Power Systems Research*, vol. 80, no. 3, pp. 256–264, 2010.

[6] J. W. Taylor, "Short-term electricity demand forecasting using double seasonal exponential smoothing," *Journal of the Operational Research Society*, vol. 54, no. 8, pp. 799–805, 2003.

[7] A. Setiawan, I. Koprinska, and V. G. Agelidis, "Very short-term electricity load demand forecasting using support vector regression," in *Proceedings of the 2009 International Joint Conference on Neural Networks*, 2009, pp. 2888–2894.

[8] A. Almalaq and G. Edwards, "A review of deep learning methods applied on load forecasting," in *2017 16th IEEE international conference on machine learning and applications (ICMLA)*. IEEE, 2017, pp. 511–516.

[9] K. Amarasinghe, D. L. Marino, and M. Manic, "Deep neural networks for energy load forecasting," in *Proceedings of the IEEE 26th International Symposium on Industrial Electronics (ISIE)*, 2017, pp. 1483–1488.

[10] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27–48, 2016.

[11] M. Tzelepi and A. Tefas, "Deep convolutional learning for content based image retrieval," *Neurocomputing*, vol. 275, pp. 2467–2478, 2018.

[12] N. Passalis, M. Tzelepi, and A. Tefas, "Heterogeneous knowledge distillation using information flow modeling," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 2339–2348.

[13] M. Tzelepi, N. Passalis, and A. Tefas, "Online subclass knowledge distillation," *Expert Systems with Applications*, vol. 181, p. 115132, 2021.

[14] M. Tzelepi and A. Tefas, "Forecasting day-ahead electric load demand on greek energy market," in *Thirteen IEEE International Conference on Information, Intelligence, Systems and Applications (IISA),*. IEEE, 2022.

[15] S. F. Ahmed *et al.*, "Short-term electrical load demand forecasting based on lstm and rnn deep neural networks," *Mathematical Problems in Engineering*, vol. 2022.

[16] D. Zhao, Q. Ge, Y. Tian, J. Cui, B. Xie, and T. Hong, "Short-term load demand forecasting through rich features based on recurrent neural networks," *IET Generation, Transmission & Distribution*, vol. 15, no. 5, pp. 927–937, 2021.

[17] P. Bunnoon, K. Chalermyanont, and C. Limsakul, "Mid-term load forecasting: Level suitably of wavelet and neural network based on factor selection," *Energy Procedia*, vol. 14, pp. 438–444, 2012.

[18] J. Hartono, A. Surya, R. Utami, B. Harsono, H. Tambunan, and A. Purnomoadi, "Long term load demand forecasting in bali province using deep learning neural network," in *2020 International Conference on Technology and Policy in Energy and Electric Power (ICT-PEP)*. IEEE, 2020, pp. 174–178.

[19] A. N. Adamakos and M. K. Titsias, "Short-term load forecasting using a cluster of neural networks for the greek energy market," in *Proceedings of the 9th Hellenic Conference on Artificial Intelligence*, 2016, pp. 1–6.

[20] N. Maragkos, M. Tzelepi, N. Passalis, A. Adamakos, and A. Tefas, "Electric load demand forecasting on greek energy market using lightweight neural networks," in *2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*. IEEE, 2022, pp. 1–5.

[21] G. Sideratos, A. Ikonomopoulos, and N. D. Hatziargyriou, "A novel fuzzy-based ensemble model for load forecasting using hybrid deep neural networks," *Electric Power Systems Research*, vol. 178, p. 106025, 2020.

[22] N. Andriopoulos, A. Magklaras, A. Birbas, A. Papalexopoulos, C. Valouxis, S. Daskalaki, M. Birbas, E. Housos, and G. P. Papaioannou, "Short term electric load forecasting based on data transformation and statistical machine learning," *Applied Sciences*, vol. 11, no. 1, p. 158, 2021.

[23] A. Rafati, M. Joorabian, and E. Mashhour, "An efficient hour-ahead electrical load forecasting method based on innovative features," *Energy*, vol. 201, p. 117511, 2020.

[24] V. Bui, T. L. Pham, J. Kim, Y. M. Jang *et al.*, "Rnn-based deep learning for one-hour ahead load forecasting," in *2020 International conference on artificial intelligence in information and communication (ICAIIC)*. IEEE, 2020, pp. 587–589.

[25] J. Silva, I. Praça, T. Pinto, and Z. Vale, "Energy consumption forecasting using ensemble learning algorithms," in *International Symposium on Distributed Computing and Artificial Intelligence*. Springer, 2019, pp. 5–13.

[26] A. Laouafi, M. Mordjaoui, and D. Dib, "One-hour ahead electric load forecasting using neuro-fuzzy system in a parallel approach," in *Computational intelligence applications in modeling and control*. Springer, 2015, pp. 95–121.

[27] H. Qiao, K. Chalermyanont, and R. Duangsoithong, "Hour-ahead power load demand time series forecasting using four methods in three cases," in *2019 16th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*. IEEE, 2019, pp. 593–596.

[28] I. E. Livieris, E. Pintelas, and P. Pintelas, "A cnn–lstm model for gold price time-series forecasting," *Neural computing and applications*, vol. 32, no. 23, pp. 17 351–17 360, 2020.

[29] X. Zhang, X. Liang, A. Zhiyuli, S. Zhang, R. Xu, and B. Wu, "At-lstm: An attention-based lstm model for financial time series prediction," in *IOP Conference Series: Materials Science and Engineering*, vol. 569, no. 5. IOP Publishing, 2019, p. 052037.

[30] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin *et al.*, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467*, 2016.

[31] S. Mehtab and J. Sen, "Stock price prediction using convolutional neural networks on a multivariate timeseries," *arXiv preprint arXiv:2001.09769*, 2020.

[32] A. Sagheer and M. Kotb, "Time series forecasting of petroleum production using deep lstm recurrent networks," *Neurocomputing*, vol. 323, pp. 203–213, 2019.

[33] C. Fan, Y. Zhang, Y. Pan, X. Li, C. Zhang, R. Yuan, D. Wu, W. Wang, J. Pei, and H. Huang, "Multi-horizon time series forecasting with temporal attention learning," in *Proceedings of the 25th ACM SIGKDD International conference on knowledge discovery & data mining*, 2019, pp. 2527–2535.