Next-day Electricity Demand Forecast: A New Ensemble Recommendation System Using Peak and Valley

Bader Alaskar*‡, Abdullah Alhadlaq*‡ Meshal Alharbi*‡, Saud Alghumayjan*‡,
Ahmad Alabdulkareem*‡, Mansour Alsaleh*‡ and Devavrat Shah†
*The Center for Complex Engineering Systems at KACST and MIT
Email: {balaskar, aaalhadlaq, masaalharbi, salghumayjan, aalabdulkareem, maalsaleh}@kacst.edu.sa
†Massachusetts Institute of Technology, Cambridge, MA, Email: devavrat@mit.edu
‡Corresponding authors

***Abstract*—Electricity demand forecast plays a major role in the planning and resource allocation phase of utility companies. In particular, predicted peak and valley (PaV) demand points seems critical, as they determine the maximum required generation capacity and baseload to meet the minimum underlying demand, respectively. In this paper, we propose multiple techniques to enhance day-ahead forecasting models by leveraging independent daily PaV predictors to ensemble short-term electricity demand forecasters. These ensemble techniques are then incorporated into a novel ensemble recommendation system (ERS). The ERS suggests the most appropriate ensemble technique to enhance the day-ahead predictor's performance while minimizing the computation required for testing multiple ensemble algorithms, relative to a single ensemble algorithm. This approach aims to improve the PaV forecasting and to enhance the overall accuracy of the day-ahead forecaster and it can be used with any combination of forecasting models. We demonstrate the effectiveness of our approach through a case study using a time-series prediction database model (tspDB) and a deep neural network (DNN) model for predicting the demand of the next day. The results show an improvement of 33% and 12% in the mean absolute percentage error of the forecasted PaV points using the tspDB and DNN models, respectively, as well as, enhancement in the overall day-ahead forecast.**

***Index Terms*—Electricity Demand Forecasting, Peak and Valley Ensemble, Ensemble Recommendation System.**

I. Introduction

For electric power system operators worldwide, predicting electricity demand is an essential procedure in the dispatching process before delivering electrical energy to the consumers. Therefore, electricity demand forecasting (EDF) has attracted significant attention in the field of time-series forecasting. For any given load profile, the optimal sequence of generator units to be committed and the energy they should dispatch are calculated using unit commitment and economic dispatch (UCED) algorithms [1, 2]. The accuracy of the optimal solution generated from the UCED algorithms depends highly on the accuracy of the forecasted demand. Overestimating or underestimating the load could cost millions of dollars per year or may compromise the grid security, which reveals the importance of having an accurate EDF model. Although recent advancements in machine-learning algorithms have enabled EDF models to achieve high accuracy levels, researchers are continually exploring new methods to enhance the predictive ability of these models (e.g., see [3, 4]), as small accuracy enhancements can lead to large savings for utility companies.

Amidst all the points in a given load profile time-series, the peak and valley (PaV) points have received special consideration in the power systems field [5, 6]. Errors in forecasting the peak point may cause more technical and economic losses than errors in other points of the day as, in general, with higher demand, more expensive and less efficient generators are used to meet the demand. Conversely, the valley point of the load profile dictates the baseload, which is an essential piece of information. The baseload specifies the energy that will be consumed at all times of the day, and the utility company can assign highly efficient generators with low ramping rates to supply this static load.

The criticality of the PaV points extends even further. If the PaV points are predicted efficiently, their predicted values can be used to enhance the accuracy of the forecasted load profile. Researchers have presented methodologies for ensembling a prediction of a day-ahead profile with a prediction of the PaV points to produce an improved approximation of the actual load profile (e.g., see [7, 8, 9]). However, the approach to select a suitable ensembling technique to maximize the enhancement of the next-day load prediction has not been examined in the literature. Further, in the reported results, the accuracy of the algorithms is determined by calculating the final forecasting error after applying the ensemble methodology. However, this error is dependent on the accuracy of the PaV prediction model, and thus it does not single out the performance of the method. Therefore, a performance evaluation method that is independent of the PaV models is required.

In this work, we examine various techniques to enhance next-day predictions using independent PaV point predictors and propose in detail a method for evaluating the performance of these techniques. The methodology can be considered as an ensemble recommendation system (ERS) that aids in the evaluation of different ensemble techniques and selects the most fitting one for the case in question.

978-1-7281-8897-3/21/$31.00 ©2021 IEEE

Contributions. Our major contributions are as follows:

1) We developed two ensemble algorithms that enhance the predictions of PaV points and the hours close to these points.
2) We proposed a methodology for measuring the performance of the ensemble algorithms irrespective of the performance of the PaV prediction models.
3) We constructed an ERS that suggests the best-suited ensemble algorithm depending on the accuracy of the PaV predictors.
4) We demonstrated the effectiveness of the proposed model on a real-world use case of forecasting the next-day electricity demand for a large utility company covering a population of over 30 million people.

Roadmap. The remainder of this paper is organized as follows: First, the dataset and models used in Section II are discussed. Section III describes in detail the ensemble techniques. Thereafter, a case study in which the proposed work was applied, and the obtained results are presented in Section IV. Section V explains the proposed ERS. Finally, Section VI summarizes the findings and provides several concluding remarks.

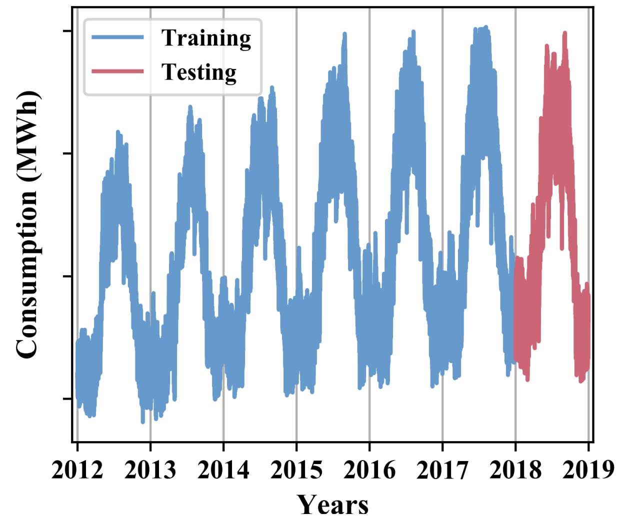

Fig. 1. Load curve used in training and testing EDF models

II. Preliminaries

Dataset. In our case study, the used time-series consisted of the electricity demand data for a large utility company covering a population of over 30 million people. The data were processed and cleaned. The cleaning procedure began by removing the outliers beyond five standard deviations from the mean. Thereafter, sudden changes in the demand that lay beyond reasonable operation were removed. Moreover, any linear interpolations and monotonic windows were removed. Finally, the removed sections of the data were imputed using a methodology proposed in our previous work [10], which uses singular value decomposition to approximate the missing parts of the time series with reasonable accuracy. The data ranged from the first day of the year 2012 until the final day of the year 2018 with an hourly resolution as shown in Figure 1, where the horizontal and vertical lines represent the time and consumption, respectively. This yielded 61,368 points in the time series, each with an electricity demand corresponding to the hour that is represented. As depicted in Figure 1, the data of the period from 2012 to 2017 (86% of all the data) were used for training, whereas the data of the year 2018 (14% of all the data) were used for testing. These data were used to train and validate the day-ahead predictors.

Used Day-Ahead Forecasting Models. To test the proposed methodology for enhancing the EDF models, two models are used to perform day-ahead forecasting, which are referred to as the next-day predictors. The first is an off-the-shelf time series prediction database (tspDB) model [11], which uses PostgreSQL and allows the user to perform predictive querying by imputing missing or corrupted data or forecasting future points in a time series. The second is the deep neural network (DNN) model, a known method that can predict patterns with high accuracy and is used in the field of electricity demand forecasting [12, 13, 14]. For that purpose, we developed an in-house trained DNN model, which consists of an input layer (that takes one week of lag values) followed by four dense layers, each consisting of 128 units. The final layer (the output layer) consists of 24 units that represent the load of the following day. The tspDB and DNN models is used to predict the demand for the next day using demand lags as features.

III. Ensemble Next-Day Forecaster using PaV Predictors Methodology

This section presents new ways to ensemble the next-day predictor with the independent PaV predictors of the next day. Figure 2 shows the block diagram that describes the ensemble methodology that the proposed ensemble algorithms follow. F_o is a vector consisting of 24 points predicted by the next-day predictor, referred herein as the original predictor. The upper and lower blocks represent respectively the PaV points, p and v, predicted by the independent PaV predictors. The two plots on the sides showcase the changes that might occur when ensembling the next-day predictor.

$$F_{m:1} = \frac{(F_o - v_o)(p - v)}{p_o - v_o} + v \tag{1}$$

The original predictor F_o is modified in Equation 1. v_o and p_o represent the valley and peak of the original predictor, respectively. p and v represent the independently predicted PaV points, respectively. As a result, $F_{m:1}$ is the modified next-day prediction, which is represented as a vector of 24 points. This equation is used by Amral et al. [7] to ensemble the next-day predictor given effective and accurate PaV predictors. This method is used for comparison and is referred to as the all horizon (AH) ensemble.

$$\Delta F = F_{m:1} - F_o \tag{2}$$

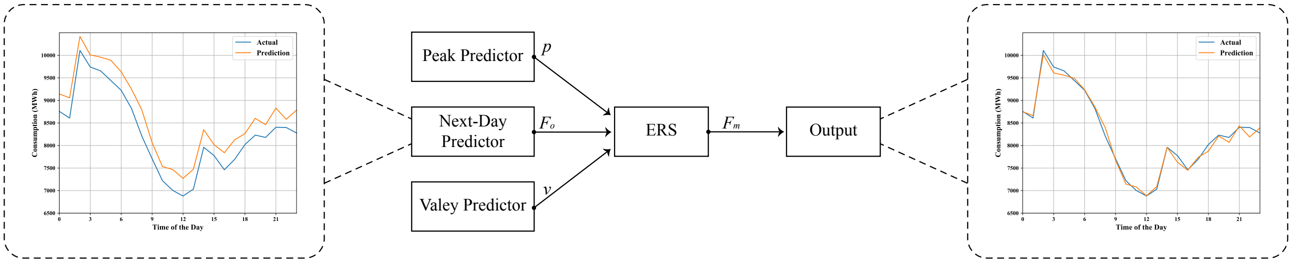

Fig. 2. Block diagram of the ensemble recommendation system methodology

The next step consists of subtracting the modified next-day predictor from the original predictor to obtain ΔF, as indicated in Equation 2.

$$F_{t,m:2} = F_{t,o} + \Delta F_t \times \left(1 - \frac{2k_t}{|indx(p) - indx(v)|}\right) \tag{3}$$

$$F_{t,m:3} = F_{t,o} + \Delta F_t \times e^{-k_t} \tag{4}$$

The two developed ensemble algorithms are illustrated in equations 3 and 4. The predicted demand is split into two regions, where the splitting point is the middle point between the PaV's occurrence time. This split ensures to have a peak in one region, and a valley on the other. Both equations depend on the time location of PaV on the predicted vector, indicated as $indx(p)$ and $indx(v)$, and $[k = 0, 1, ..., \frac{|indx(p)-indx(v)|}{2}]$ is the absolute time distance of the predicted point from the peak time, if the point lies on the peak region, or valley time, in case the point falls in the valley region. Equation 3 applies linearly decaying modification on each instance t. Based on the region, where the instance fall (either peak or valley region), the absolute time distance k_t is measured. Applying this method on the predicted points by the original predictor F_o results in less modification as the point is further from the PaV location. This ensemble technique will be named the linear decay (LD) ensemble. Equation 4 has an exponential decay function that significantly modifies the PaV points and the points that are closer thereto, whereas it does not affect the points further from the PaV in the original predictor. This technique will be referred to as the exponential decay (ED) ensemble. The incorporation of AH, LD, and ED ensembles will be explained in Section V.

IV. Experiment

In this section, a case study is presented to implement the proposed ensemble algorithms presented in Section III on the next-day predictors using independent PaV predictors, and the results are reported. The used electricity consumption data and how it is split between training and testing are described in Section II. Two next-day predictors were used: the in-house developed DNN and off-the-shelf tspDB, to forecast each hour of the next day for 2018, as discussed in Section II. Furthermore, as the DNN model is commonly used in time-series prediction [12], an independent DNN model was developed to forecast the PaV points of the next day for the same year of testing. The historical consumption data are used to train the PaV predictors. The features used are hourly week lags and daily month lags of peaks for the peak predictor and valleys for the valley predictor. In the following section, a comparison of performance in the prediction of PaV points between the next-day and PaV predictors will be conducted.

A. PaV Point Prediction

This section evaluates the PaV predictions from the original next-day predictors (DNN and tspDB) and the independent PaV predictor models. The performance of PaV predictions is displayed in Figure 3. Each model has left and right bars, which represent the mean absolute percentage error (MAPE) of peak and valley points, respectively. The independent PaV models are trained for predicting the PaV of the next day, and it compares the MAPE with that of the PaV point prediction from the original next-day predictors. The independent PaV predictors outperform the original next-day predictors in terms of PaV points prediction. These improved predictions of PaV points will be used to ensemble the original next-day predictors. The results of ensembling the next-day predictor will be presented in the next subsection.

B. PaV Ensemble Algorithms

This section describes the performance of the original next-day predictors, before and after applying the PaV ensemble algorithms discussed in Section III. The used PaV points in the ensemble algorithm are taken from the independent PaV predictors. Table I presents the average MAPEs of the next-day predictors for the year 2018, before and after applying all of the ensemble methods. Depending on the accuracy of the original next-day predictor, some of the ensemble algorithms might decrease the accuracy while others improve it. Looking at the tspDB next-day predictor, the best performing model in terms of MAPE being highlighted is the AH ensemble, which shows a 14.8% enhancement in the average MAPE. Moreover, a slight enhancement of the DNN predictor's average MAPE, in addition to the 12% improvement in predicting PaV points, is obtained using the proposed ED ensemble.

Multiple ways to ensemble the following day predictors are proposed. The next section will address a novel way to validate these algorithms and suggest the most suitable ensemble algorithm that would lead to the maximum improvement to the next-day predictor.

V. PaV Ensemble Recommendation System (ERS)

The previous section validated the developed ensemble algorithms. This section will propose the ERS methodology,

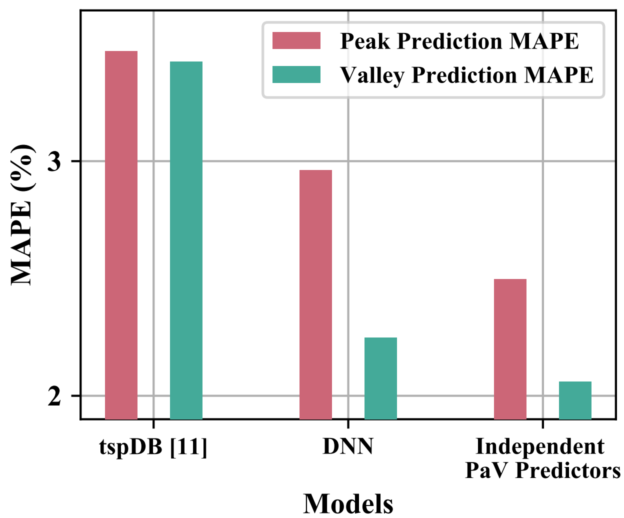

Fig. 3. MAPEs of PaV points for all models

displayed in the block diagram in Figure 2, which helps to guide the selection of the ensemble algorithms based on the PaV predictors' performance. To begin, the construction of the ERS requires an offline validation for the ensemble algorithms based on different performances for the PaV predictors, which is simulated by adding Gaussian noise to the actual data. This offline validation helps to identify the enhancement of each ensemble algorithm for the simulated PaV performance. After the offline validation, ERS is ready to be tested online, which takes the independent PaV predictors performance as input, and outputs the best ensemble algorithm to be applied to the original predictor F_o.

To start the ERS's offline validation, the next-day predictors (described in Section II) are used to generate the prediction of the next day F_o. Next, to sweep across all the possible performances for the independent PaV predictors, the actual PaV points of the next day are chosen as the initial state. After that, the chosen PaV points are used to ensemble the next-day predictors using all of the described ensemble algorithms. Thereafter, the MAPE of the next-day predictor is calculated for all the ensemble methods.

$$MAPE(\%) = \frac{100}{24 \times 365} \sum_{j=1}^{365} \sum_{t=1}^{24} \left| \frac{A_{t,j} - F_{m,t,j}}{A_{t,j}} \right| \tag{5}$$

Equation 5 is used to calculate the MAPE across the testing year (2018) for each ensemble technique applied to the next-day predictor F_o. $A_{t,j}$ represents actual demand for the point at time t at day j, and $[F_{m,t,j}\, where: \; m = [1,2,3]]$ denotes all the ensemble methods predicted at time t for day j.

$$p = p_a + z, \qquad where: \; z \sim N(0, \sigma) \tag{6}$$

$$v = v_a + z, \qquad where: \; z \sim N(0, \sigma) \tag{7}$$

After measuring the MAPE, Gaussian noise is added to the chosen PaV points. Equations 6 and 7 refer to the peak and valley points, respectively, where z is independent and identically distributed with a zero-mean normal distribution

TABLE I

COMPARISON OF AVERAGE MAPES BETWEEN THE ORIGINAL NEXT-DAY PREDICTORS WITH DIFFERENT ENSEMBLE TECHNIQUES

	DNN	tspDB
Original	2.74	3.78
AH ensemble	2.82	3.22
LD ensemble	2.73	3.39
ED ensemble	2.72	3.6

and with a variance of σ (the noise). The noise is added to simulate various independent PaV predictors' performance. Repeating the process by adding the noise to actual PaV and applying Equation 5 would results in mapping the PaV's performance with the original predictor's MAPE, as illustrated in Figures 4 and 5.

As shown in Figure 4, the validation metric applied to the next-day predictor DNN model. The y-axis shows the MAPE of the modified predictor for all the ensemble algorithms. Owing to visualization constraints, the x-axis represents the MAPE of PaV predictors, where the MAPE of independent PaV predictors is assumed to be the same, and the zero MAPE represents the actual PaV of the next day. The horizontal red dashed line represents the DNN predictor's MAPE without using any ensemble method. The validation metric declares that for a certain MAPE of PaV predictors, the MAPE of the ensemble algorithms is measured for the DNN model. Each ensemble technique has a hashed region where it performs the best, and by identifying these regions the model would suggest a suitable ensemble algorithm based on the MAPE of PaV point predictors. The same process is conducted for the other next-day predictor, the tspDB model, and the regions are determined as shown in Figure 5.

The suggested criteria are reliant on the performance of the independent PaV predictors. If the MAPE of the PaV point predictors is very high, there will not be any region where the ensemble technique would enhance the next-day predictor. In this case, the ERS would suggest not to use any ensemble method. Moreover, the system determines the best algorithm to use based on the performance of the PaV predictor inputted to the ERS. At the same time, the system cuts the number of ensemble algorithms needed to be applied in the online testing.

VI. CONCLUDING REMARKS

A methodology for enhancing the accuracy of next-day prediction algorithms has been presented in this paper. The approach consists of constructing models solely to predict the PaV points. Subsequently, depending on the MAPE of the trained PaV point predictors, an ensemble algorithm is recommended based on an ERS. The ERS will recommend the most suitable ensemble algorithm, which leads to maximum enhancement of the next-day predictor. The results demonstrated a significant decrease in the MAPE of the off-the-shelf model for the time series prediction, along with a slight enhancement in the MAPEs of the in-house developed DNN model. At the same time, significant improvement was

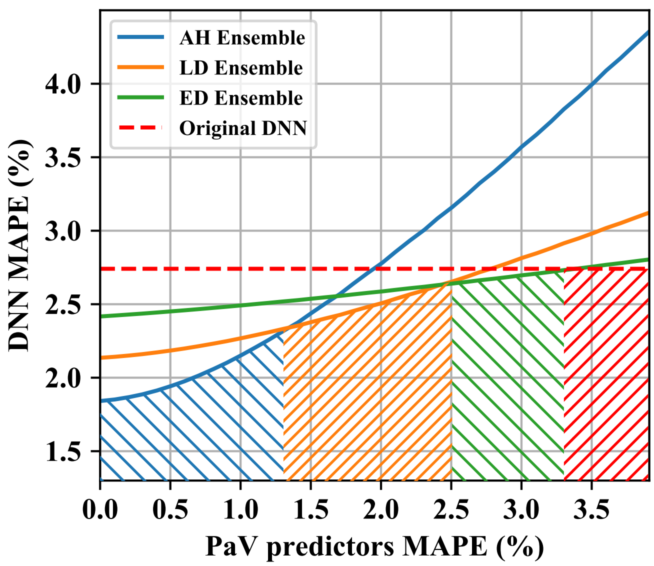

Fig. 4. MAPE of PaV predictors vs. average MAPE of DNN predictor for all the ensemble algorithms

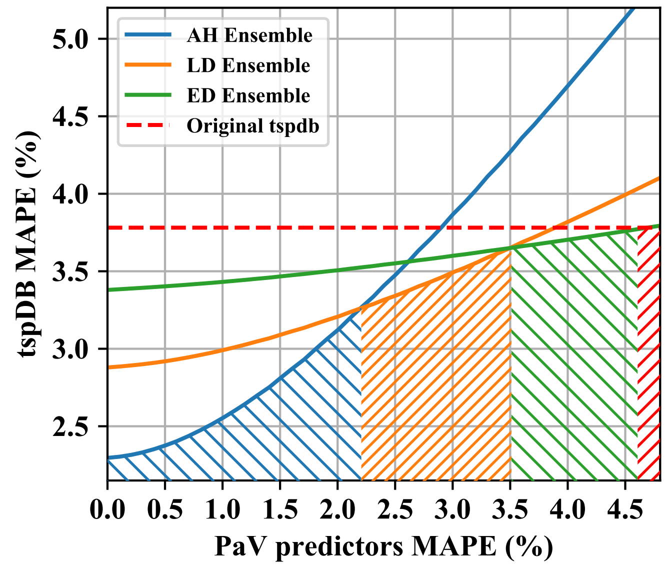

Fig. 5. MAPE of PaV predictors vs. average MAPE of off-the-shelf tspDB predictor for all the ensemble algorithms

achieved in predicting the peak and valley points, which is considered to be critical in the planning of power systems.

Although the presented method recommends the suitable ensemble technique that leads to a significant improvement in the MAPE of the next-day predictors, as demonstrated in Section IV, new ensemble techniques could be developed. Such techniques could incorporate additional time series measures, such as the mean and standard deviation, to enhance the time series predictor. The focus of this study was on the peak and valley points because of the significance of these data points to the physical power system, but an extension to the presented work could incorporate the above-mentioned measures.

Further, future potential areas include examining the performance and reliability of this methodology with time-series other than that of electricity demand. Moreover, the correction algorithms are not limited to the ones presented in this paper. It would also be interesting to investigate the reason why many off-the-shelf models appear to miscalculate the magnitudinal aspect of the demand. This may suggest alternative means of correcting the performance of the models that target the root of the issue causing the miscalculation in the first place.

References

[1] Antonio J Conejo and Luis Baringo, "Unit commitment and economic dispatch," in *Power System Operations*, pp. 197–232. Springer, 2018.

[2] R Ponciroli, NE Stauff, J Ramsey, F Ganda, and RB Vilim, "An improved genetic algorithm approach to the unit commitment/economic dispatch problem," *IEEE Transactions on Power Systems*, 2020.

[3] AS Khwaja, A Anpalagan, M Naeem, and B Venkatesh, "Joint bagged-boosted artificial neural networks: Using ensemble machine learning to improve short-term electricity load forecasting," *Electric Power Systems Research*, vol. 179, pp. 106080, 2020.

[4] Nasir Ayub, Nadeem Javaid, Sana Mujeeb, Maheen Zahid, Wazir Zada Khan, and Muhammad Umar Khattak, "Electricity load forecasting in smart grids using support vector machine," in *International Conference on Advanced Information Networking and Applications*. Springer, 2019, pp. 1–13.

[5] PK Dash, AC Liew, and S Rahman, "Peak load forecasting using a fuzzy neural network," *Electric Power Systems Research*, vol. 32, no. 1, pp. 19–23, 1995.

[6] Santisouk Phimphachan, K Chamnongthai, Pinit Kumhom, N Jittiwarangkul, and A Sangswang, "Energy and peak load forecast models using neural network for fast developing area," in *IEEE International Symposium on Communications and Information Technology, 2004. ISCIT 2004*. IEEE, 2004, vol. 1, pp. 389–393.

[7] N Amral, D King, and CS Ozveren, "Application of artificial neural network for short term load forecasting," in *2008 43rd International Universities Power Engineering Conference*. IEEE, 2008, pp. 1–5.

[8] Luis Hernández, Carlos Baladrón, Javier M Aguiar, et al., "Improved short-term load forecasting based on two-stage predictions with artificial neural networks in a microgrid environment," *Energies*, vol. 6, no. 9, pp. 4489–4507, 2013.

[9] Amit Jain and B Satish, "Integrated approach for short term load forecasting using svm and ann," in *TENCON 2008-2008 IEEE Region 10 Conference*. IEEE, 2008, pp. 1–6.

[10] Abdullah Alomar, Arwa Alanqary, Mansour Alsaleh, Anas Alfaris, and Devavrat Shah, "Adaptive approach for time series imputation using matrix estimation methods," *Submitted to the Pattern Recognition Letters*.

[11] Anish Agarwal, Abdullah Alomar, Muhammad J Amjad, Robert Lindland, and Devavrat Shah, "tspdb: Time series predict db," *arXiv preprint arXiv:1903.07097*, 2019.

[12] Stefan Hosein and Patrick Hosein, "Load forecasting using deep neural networks," in *2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 2017, pp. 1–5.

[13] Tareq Hossen, Siby Jose Plathottam, Radha Krishnan Angamuthu, Prakash Ranganathan, and Hossein Salehfar, "Short-term load forecasting using deep neural networks (dnn)," in *2017 North American Power Symposium (NAPS)*. IEEE, 2017, pp. 1–6.

[14] Pyae Pyae Phyo and Chawalit Jeenanunta, "Electricity load forecasting using a deep neural network," *Engineering and Applied Science Research*, vol. 46, no. 1, pp. 10–17, 2019.