2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)

Ensemble learning models for short-term electricity demand forecasting

Ahmed Ghareeb
College of Engineering, University of Kirkuk
aghareeb@uokirkuk.edu.iq

Hussein Al-bayaty
College of Engineering, University of Kirkuk
dr.hussein@uokirkuk.edu.iq

Qubad Haseeb
College of Engineering, University of Kirkuk
qubadsabah@uokirkuk.edu.iq

Mohammed Zeinalabideen
College of Engineering, University of Kirkuk
mohammed.jawad@uokirkuk.edu.iq

Abstract— Forecasting energy consumption is critical in decision-making for efficient energy saving, improve stability of the power grid, and prevent supply-demand discrepancy. To predict day- ahead load forecasting for the demand of city of Kirkuk two scenarios were presented. First, benchmarked three individual machine learning algorithms e.g. generalized linear model (GLM), artificial neural network (ANN), and random forest (RF). Second, compared the predictive capabilities for individual models with the ensemble models. The results indicate that the predictive models maybe can be improved using simple ensemble learning strategies such as averaging the predicted results. This study is also present future research directions to improve the model prediction capabilities.

Keywords— Energy informatics, Load forecasting, Artificial neural network, Random forest, Generalized linear model.

I. INTRODUCTION

Electricity is regarded as the main component of national and international growth in the worldwide. The increased demand for electricity increases manufacturing needs and the use of fossil fuels that have already caused air pollution; and harmful effects on human health and the global atmosphere including acidic rain caused by global warming; and other sources of contaminations [1] .

In Machine Learning (ML), ensemble methods merge different learners to analyze forecasting based on essential learning techniques [2]. The typical Ensemble Learning (EL) methods contain gathering of bootstrap (or bagging) and boosting techniques. Random Forest (RF); in example, bagging merges random decision trees and can be used for regression, classification, and other duties. The validation and influence of (RF) for regression has been examined and analyzed in [3]. Another method called boosting method uses to emphasize misclassified cases by adding new examples which results in building an ensemble model in order to earn a competitive performance for time series prediction [4-5]. As the most common use of boosting, Ada- Boost [6] was assessed with other ML algorithms such as support vector machines (SVM) [7] and merged with this algorithm to further improve for prediction attitude [8]. In addition, stacking [9] is an example of (EL) multiple algorithms, this method mixes the results which are created by different base learners in the first level. Furthermore, it aims to incorporate the results of these simple learners in an optimal way to improve the capacity to generalize [10]. Multistep predictions have harder jobs than the one-step predictor, owing to absence of information and collection of errors. Therefore, multistep predictions are preferred in different applications. Recently, various methods of forecasting were recommended to solve some genuine problems as a globally forecasting challenge [11]. In many studies, researchers compared the long-term prediction execution of the hybrid model [12]. On the other hand, Ardakani et al. [13] suggested new model called optimal artificial neural networks (ANN) model relied on the amended version of particle swarm optimization in order to determine long-term electrical energy consumption. Regarding the same aspect this study, [14] introduced a model named the hybrid-connected complex neural network (HCNN), which is able to capture the dynamics embedded in chaotic time series and to predict long horizons of such series. In [15], researchers combined models with self-organizing maps for long-term forecasting of chaotic time series.

At the same time, the authors in [14] presented a model called the hybrid-connected complex neural network (HCNN), this model can predict long distances of chaotic time series and has the ability to latch the dynamics embedded in a disordered time series. Also, the authors in [15], merged models with self-organizing charts for unpredictable time series long term forecasting.

Moreover, models like ANN and SVM used in short-term forecasting modelling, performed greatly for the job of one-step prediction[16-17].Nevertheless, the attitude is changed when using with the general multistep issues as these models suffer regression and difficulties.

In general, the long-term forecasting models act well in multistep prediction; because these types of models are constructed for long time forecasting tasks (could be weekly or monthly time series prediction), whereas one-step ahead prediction is worse than other techniques. Commonly, combined prediction models (for instance, mixing both short-term and long-term methods) give better results compared to single models [18], due to ability to combine the advantages of both approaches (long and short terms) and at the same time avoid their drawbacks.

Likewise, dynamic sets approaches[19–21] such as bagging or boosting examined to merge the outcomes of complementary and various models produced by resampling training data, reweighting, and active perturbing, however,

2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI) | 978-1-7281-9675-6/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICDABI51230.2020.9325623

main static combination methods such [22] and [23] depend on allocate a constant weight for each model like (inverse mean and average). In a result, a construction dependent weights can be utilized in preventing the disadvantages of a static and dynamic collection such as short and long-term forecasts [24].

Many factors influence the precision and accuracy of predicting electricity consumption. In electric load forecasting, it was verified that the predictability depends not only on the computational efficiency of the algorithms used, but also on the consistency of the data analyzed [25] and the ability to integrate significant exogenous factors in the model [24]. Accordingly, models have been improved by capturing the association between the factors of load and weather variables including humidity and temperature, and it is implemented by adding some key factors such as lagged hourly temperatures and/or daily moving average temperatures [26].

Demand forecasts may be required for various time horizons: long, medium and short term. Short term load forecasting (STLF) is the system load prediction on an hourly basis, with an interval between 1 hour and 1 week, and the methodologies in STLF are classified by Huang into: Regression techniques, Time series, State estimation, Exponential smoothing [27], Expert (intelligent) system, and Artificial Neural Networks (ANN) [28] [29]. Another classification presented by Chen and Lie [30] for energy-demand temperature models comprises: bottom up demand models and top-down demand models. Wherein Bottom-up models require specific collection of data on meteorological, demographic, and energy usage for each sector to model the relationship between temperature and demand, and Top-down demand model uses regression equations to forecast energy demand from climate variables such as temperature, precipitation etc. [31].

In recent years, several studies in Iraq have attempted to study and analyze the possibilities for demanding electrical energy for many forecasting horizons, e.g. short term and long term and perspectives [32]. Cankurt et. al [33] proposed an artificial natural network and sliding window techniques to examine daily forecasts of maximum energy demand in northern Iraqi region. Other papers for predicting energy demand proposed by many researchers include time series analysis using ARIMA [34] and SARIMA [35] [36] models. Moreover, in ANN, Recurrent Neural Network architecture [37] where used more frequently e.g. Elman neural network[38] [39]. The literature also include application of electricity generation from waste, e.g. Municipal solid waste (MSW) [40], spatial distribution of shared electrical generators [41], the factors influencing consumer behavior in the electric appliances market [42] and home energy management controller (HEMC) considering demand-response (DR) events [43]. Ultimately, comparative study for forecasting electricity demand proposed in [44], and Hybrid models in [45] and [46].

Ensemble learning was applied in the early 1990s as a more comprehensive data mining methodology [47 - 55]. Combined multiple predictions is a consolidated strategy for enhancing predictive accuracy. Marziali et. al [47] forecasted short-term forecasting of gas demand using ensemble learning, and suggested it might lead to significant predictive model performance. Barrow et. al [48], showed the effectiveness of combining forecasting models in reducing the shape and distribution of forecasting errors. Moreover, other related papers in the literature include using ensemble learning in various frameworks shown in [49]. This paper represents an update and improvement to our previously published article shown in [50] for the city of Kirkuk demand forecasting using machine learning techniques. The goal of this study is to investigate the accuracy of each predictive model and explore the possibility of model combinations in improving prediction precision. The performance by improving predictive accuracy, using ensemble learning is also discussed. In the next section, the first segment introduces the data as well as giving insights about data preparation for prediction, and the second segment includes a description about the candidate algorithms selected in this study.

II. Data Analysis Methodologies

A. *Data Description and Prediction*

The data comes from hourly collection of loads by the National Dispatch Control Center (NDCC) for the Kirkuk governorate. Since the paper publication [50], additional data were collected and the datasets updated to include two years and three months of electricity demand for the city of Kirkuk e.g. 2018-2019 and three months from 2020. To make fare comparison between this model and our previously published work, models benchmarked using one year of data each time. In other words, the models were trained and compared to our work in 2018 (using one year of data), then validated again by training the next year (2019) and made prediction for out-of-sample data of 3 months in 2020. In this paper three algorithms were selected: Artificial Neural Network (ANN), Random Forest (RF), and Generalized Linear Model (GLM). The algorithms implemented using R-programming software.

B. *Algorithms*

This study uses quantitative statistical techniques for predicting electricity demand, and it tends to be quite tied to the consideration of prediction precision and data usage. Three algorithms selected, the first one is the artificial neural network (ANN). ANN is essentially a human brain mathematical representation consisting of numerous neurons linked to each other. It has imitation ability as in the human brain to take decisions and proven adaptive performance for a complicated and noisy set of information. The architectures of the ANN model presented in this paper are feeding-forward neural networks, and the configuration includes a single hidden layer, and multinomial log-linear models and 10 hidden nodes.

The second algorithm is Random Forest (RF), which was introduced by Breiman in 2001. They may be used for classification tasks in which classification trees are the base models, or regression tasks where regression trees are the base models. Random forests are a way to average deep decision trees, trained on specific segments of the same training set, with the objective of minimizing variance. This comes at the cost of a slight increase in bias and some lack of interpretability, but usually dramatically improves efficiency in the final model. The random forest training algorithm applies the general strategy of bootstrap averaging, or bagging, to tree learners. This leads to improved model results because it eliminates the variability of the model without increasing the bias. The number of trees to build usually by default in R-programming software selected to be 500 as a good resulting model. The number of the variables to be considered at any time in deciding how to partition the dataset in random forest algorithm is selected to be 10.

The third algorithm is Generalized Linear Model (GLM). It is a basic version of linear regression in which an error distribution rather than a regular distribution is permitted for the response variable. Moreover, it generalizes a straightforward linear regression by allowing the model to be connected to the response variable by a relation function which permits the variance of each sample to be a function of its predicted value [55].

III. RESULTS ASSESSMENT

In terms of tests to validate the predictive accuracy of our candidate models, the mean absolute percentage error (MAPE) is used; for forecasting performance this evaluation metric is been used as an accuracy measurement for all of the created models to estimating the error rate shown below:

$$\text{MAPE} = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \quad (1)$$

F

Where A_t is the actual value and F_t is the forecast value.

$$R^2 \equiv 1 - \frac{SS_{res}}{SS_{tot}} \quad (2)$$

Coefficient of determination (R^2) estimates the percentage of variability in the dependent variable described by all of the model's independent variables. Predicted R-squared to evaluate estimates about how well a regression model performs. Figures (1-3) helps us to distinguish situations where the model is a good match for current data but it does not imply the goodness at forecasting. For an estimation of the success of the models on new data, it is necessary to apply it to data that is unseen. The amount of unseen data in this paper is the first three months of 2020.

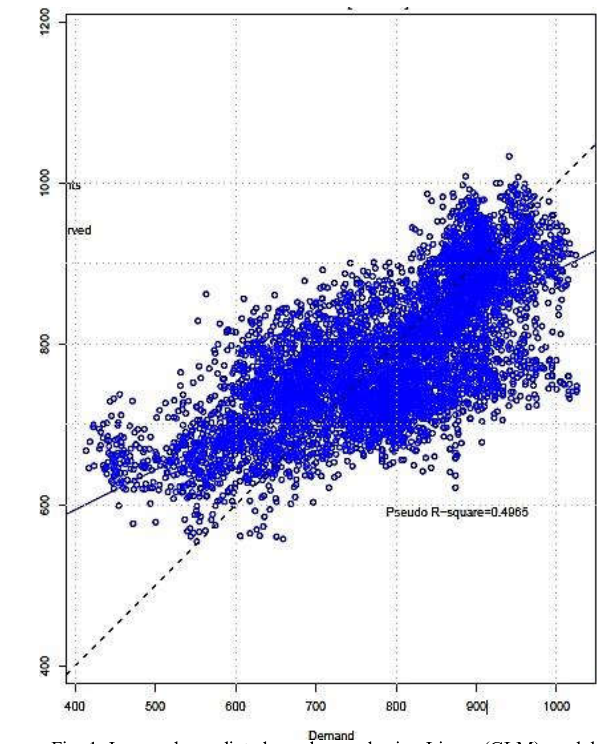

Fig. 1: In-sample predicted vs. observed using Linear (GLM) model

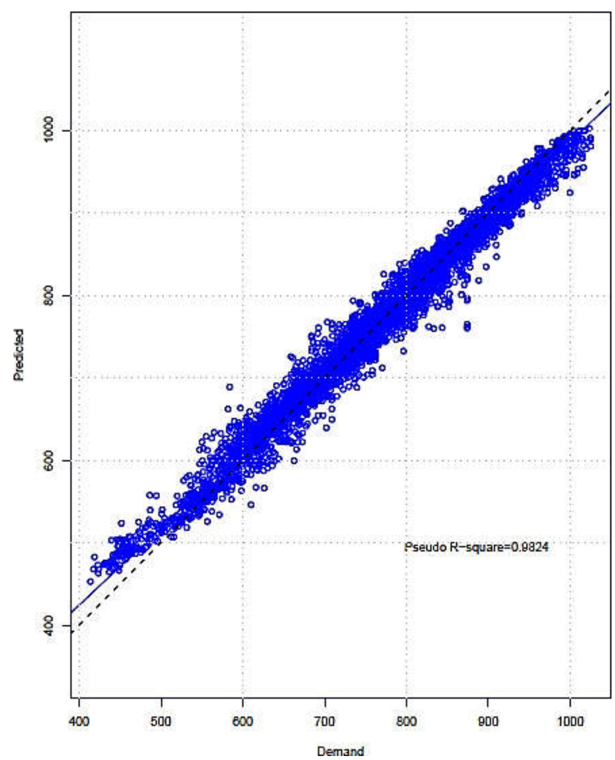

Fig. 2: In-sample predicted vs. observed using Random Forest (RF) model

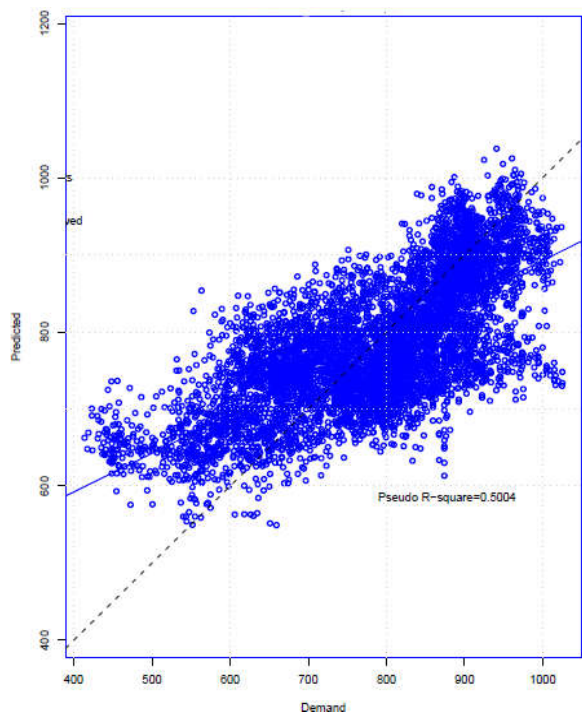

Fig. 3: In-sample predicted vs. observed using Artificial Neural Network (ANN) model

The in-sample Pseudo R-squared values obtained for GLM, RF, and ANN are approximately: 0.49, 0.98, and 0.50 respectively. The below tables (1,2) show the prediction accuracy in average for applying each developed model to three months of out-of-sample data in 2020 to predict day-ahead electricity demand for the city of Kirkuk. In table 2, the predicted values were combined by simply averaging multiple outcomes of the predictive models.

Table I: Results of Using individual predictive models

Evaluation metrics	GLM	ANN	RF
MAPE	4.67	4.62	3.98

Table II: Results of model combination scenarios by simple averaging the predictive models

Evaluation metrics	(GLM+ANN +RF) models	GLM +ANN	GLM +RF	ANN+RF
MAPE	4.25	4.54	4.14	4.12

Comparing the two tables, it can be seen from the results the least error is for out-of-sample prediction is for Random Forest (RF) model with MAPE of 3.98%. However, this result may be also concluding the error rates can be decreased when it is combined. Eventually, this analysis also show discrepancy in results of in-sample and out-of-sample prediction. It is clear that the model is only capable of predicting up to 24 hours with accurate precision for this dataset.

IV. Conclusion

Using individual predictive models, the results suggest that Random forest is outperforming other algorithms. However, the overall results comparing the two tables for this dataset may also indicate as the opportunity of improving the predicted models using ensemble learning strategies such as simple averaging. This study also shows a disparity in sample and out-of-sample prediction results. Further investigation for these outcomes may include benchmarking by manipulating the predictive models, adding new candidate variables, combining more than three algorithms, and voting and weighted average ensembles.

References

[1] M. T. Chaichan and H. A. Kazem, *Generating electricity using photovoltaic solar plants in Iraq*. 2018.

[2] M. R. G. Al-Shakarchi and M. M. Ghulaim, "Short-term load forecasting for Baghdad electricity region," *Electr. Mach. Power Syst.*, 2000, doi: 10.1080/073135600268315.

[3] S. R. Huang, "Short-term load forecasting using threshold autoregressive models," *IEE Proc. Gener. Transm. Distrib.*, 1997, doi: 10.1049/ip-gtd:19971144.

[4] J. G. De Gooijer and R. J. Hyndman, "25 years of time series forecasting," *Int. J. Forecast.*, 2006, doi: 10.1016/j.ijforecast.2006.01.001.

[5] J. G. De Gooijer and R. J. Hyndman, "25 Years of IIF Time Series Forecasting: A Selective Review," *SSRN Electron. J.*, 2011, doi: 10.2139/ssrn.748904.

[6] S. N. Chandramowli and F. A. Felder, "Impact of climate change on electricity systems and markets - A review of models and forecasts," *Sustain. Energy Technol. Assessments*, 2014, doi: 10.1016/j.seta.2013.11.003.

[7] H. Nie, G. Liu, X. Liu, and Y. Wang, "Hybrid of ARIMA and SVMs for short-term load forecasting," in *Energy Procedia*, 2012, doi: 10.1016/j.egypro.2012.01.229.

[8] D. K. Barrow and N. Kourentzes, "Distributions of forecasting errors of forecast combinations: Implications for inventory management," *Int. J. Prod. Econ.*, 2016, doi: 10.1016/j.ijpe.2016.03.017.

[9] T. Hong, M. Gui, M. E. Baran, and H. L. Willis, "Modeling and forecasting hourly electric load by multiple linear regression with interactions," in *IEEE PES General Meeting, PES 2010*, 2010, doi: 10.1109/PES.2010.5589959.

[10] T. Hong, J. Wilson, J. X.-I. T. on S. Grid, and undefined 2013, "Long term probabilistic load forecasting and normalization with hourly information," *ieeexplore.ieee.org*.

[11] A. Ghareeb, W. Wang, and K. Hallinan, "Data-driven modelling for building energy prediction using regression-based analysis," in *ACM International Conference Proceeding Series*, 2019, doi: 10.1145/3368691.3368701.

[12] J. Zhang, Y. M. Wei, D. Li, Z. Tan, and J. Zhou, "Short term electricity load forecasting using a hybrid model," *Energy*, 2018, doi: 10.1016/j.energy.2018.06.012.

[13] F. J. Ardakani and M. M. Ardehali, "Long-term electrical energy consumption forecasting for developing and developed economies based on different optimized models and historical data types," *Energy*, 2014, doi: 10.1016/j.energy.2013.12.031.

[14] G. Seni and J. F. Elder, "Ensemble Methods in Data Mining: Improving Accuracy Through Combining Predictions," *Synth. Lect. Data Min. Knowl. Discov.*, 2010, doi: 10.2200/s00240ed1v01y200912dmk002.

[15] U. Grömping, "Variable importance assessment in regression: Linear regression versus random forest," *Am. Stat.*, 2009, doi: 10.1198/tast.2009.08199.

[16] A. M. Youssef, H. R. Pourghasemi, Z. S. Pourtaghi, and M. M. Al-Katheeri, "Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia," *Landslides*, 2016, doi: 10.1007/s10346-015-0614-1.

[17] D. L. Shrestha and D. P. Solomatine, "Experiments with AdaBoost.RT, an improved boosting scheme for regression," *Neural Comput.*, 2006, doi: 10.1162/neco.2006.18.7.1678.

[18] J. H. Morra, Z. Tu, L. G. Apostolova, A. E. Green, A. W. Toga, and P. M. Thompson, "Comparison of adaboost and support vector machines for detecting alzheimer's disease through automated hippocampal segmentation," *IEEE Trans. Med. Imaging*, 2010, doi: 10.1109/TMI.2009.2021941.

[19] L. Guo, P. S. Ge, M. H. Zhang, L. H. Li, and Y. B. Zhao, "Pedestrian detection for intelligent transportation systems combining AdaBoost algorithm and support vector machine," *Expert Syst. Appl.*, 2012, doi:

10.1016/j.eswa.2011.09.106.

[20] R. Aldave and J.-P. Dussault, “Systematic Ensemble Learning for Regression,” *arxiv.org*, 2014.

[21] H. Al-Bayaty, M. Ambroze, and M. Z. Ahmed, “Feeding loads via harmonics utilization in AC circuit systems,” in *2016 International Conference for Students on Applied Engineering, ICSAE 2016*, 2017, doi: 10.1109/ICSAE.2016.7810192.

[22] C. Lemke and B. Gabrys, “Meta-learning for time series forecasting and forecast combination,” *Neurocomputing*, 2010, doi: 10.1016/j.neucom.2009.09.020.

[23] S. F. Crone, M. Hibon, and K. Nikolopoulos, “Advances in forecasting with neural networks? Empirical evidence from the NN3 competition on time series prediction,” *Int. J. Forecast.*, 2011, doi: 10.1016/j.ijforecast.2011.04.001.

[24] F. J. Ardakani and M. M. Ardehali, “Novel effects of demand side management data on accuracy of electrical energy consumption modeling and long-term forecasting,” *Energy Convers. Manag.*, vol. 78, pp. 745–752, 2014, doi: 10.1016/j.enconman.2013.11.019.

[25] P. Gómez-Gil, J. M. Ramírez-Cortes, S. E. Pomares Hernández, and V. Alarcón-Aquino, “A neural network scheme for long-term forecasting of chaotic time series,” *Neural Process. Lett.*, 2011, doi: 10.1007/s11063-011-9174-0.

[26] R. Fonseca-Delgado and P. Gomez-Gil, “Selecting and combining models with self-organizing maps for long-term forecasting of chaotic time series,” in *Proceedings of the International Joint Conference on Neural Networks*, 2014, pp. 2616–2623, doi: 10.1109/IJCNN.2014.6889454.

[27] J. W. Taylor, “Short-term electricity demand forecasting using double seasonal exponential smoothing,” *J. Oper. Res. Soc.*, vol. 54, no. 8, pp. 799–805, 2003, doi: 10.1057/palgrave.jors.2601589.

[28] L. F. Simmons, “Time-series decomposition using the sinusoidal model,” *Int. J. Forecast.*, vol. 6, no. 4, pp. 485–495, 1990, doi: 10.1016/0169-2070(90)90025-7.

[29] A. Abdoos, M. Hemmati, and A. A. Abdoos, “Short term load forecasting using a hybrid intelligent method,” *Knowledge-Based Syst.*, vol. 76, pp. 139–147, 2015, doi: 10.1016/j.knosys.2014.12.008.

[30] G. H. Chen and T. T. Lie, “The impact of climate change on New Zealand’s electricity demand,” in *2010 IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems, PMAPS 2010*, 2010, doi: 10.1109/PMAPS.2010.5528429.

[31] J. H. Stock and M. W. Watson, “Combination forecasts of output growth in a seven-country data set,” *J. Forecast.*, vol. 23, no. 6, pp. 405–430, Sep. 2004, doi: 10.1002/for.928.

[32] N. A. Mohammed, “Modelling of unsuppressed electrical demand forecasting in Iraq for long term,” *Energy*, 2018, doi: 10.1016/j.energy.2018.08.030.

[33] S. Cankurt and M. Yasin, “Modelling Energy Demand Forecasting Using Neural Networks with Univariate Time Series,” 2018, doi: 10.23918/iec2018.26.

[34] B. Ahmed, S. Rahim, … B. M.-Q. Z., and U. 2020, “Comparison Between ARIMA And Fourier ARIMA Model To Forecast The Demand Of Electricity In Sulaimani Governorate,” *journal.lfu.edu.krd*.

[35] Y. H. Kareem and A. R. Majeed, “Monthly peak-load demand forecasting for sulaimany governorate using SARIMA,” in *2006 IEEE PES Transmission and Distribution Conference and Exposition: Latin America, TDC’06*, 2006, doi: 10.1109/TDCLA.2006.311383.

[36] T. A. Chawsheen and M. Broom, “Seasonal time-series modeling and forecasting of monthly mean temperature for decision making in the Kurdistan Region of Iraq,” *J. Stat. Theory Pract.*, 2017, doi: 10.1080/15598608.2017.1292484.

[37] M. A. Ulkareem, W. A. Awadh, and A. S. Alasady, “A comparative study to obtain an adequate model in prediction of electricity requirements for a given future period,” in *International Iraqi Conference on Engineering Technology and its Applications, IICETA 2018*, 2018, doi: 10.1109/IICETA.2018.8458079.

[38] L. Ismael and A. I. Ismael, “Prediction the data consumption for power demands by Elman neural network,” *Int. J. Electr. Comput. Eng.*, 2019, doi: 10.11591/ijece.v9i5.pp4003-4009.

[39] H. M. Dawood, “Computational Intelligence Techniques of Long-Term Electric Power Load Forecasting on Iraqi National Grid,” *pdfs.semanticscholar.org*.

[40] O. M. Anssari, E. A. Alkaldy, N. Almudhaffar, A. N. AlTaee, and N. S. Ali, “A feasibility study of electrical energy generation from municipal solid waste in Iraq: Najaf case study,” *Int. J. Electr. Comput. Eng.*, vol. 10, no. 4, pp. 3403–3411, 2020, doi: 10.11591/ijece.v10i4.pp3403-3411.

[41] A. A. Al-waeely, S. D. Salman, W. K. Abdol-reza, M. T. Chaichan, A. Hussain, and H. S. S. Al-jibori, “Evaluation of the spatial distribution of shared electrical generators and their environmental effects : Case Study in Baghdad-Iraq,” *Int. J. Eng. Technol.*, vol. 14, no. 02, pp. 16–23, 2014.

[42] F. Furaiji, M. Łatuszyńska, and A. Wawrzyniak, “An empirical study of the factors influencing consumer behaviour in the electric appliances market,” *Contemp. Econ.*, 2012, doi: 10.5709/ce.1897-9254.52.

[43] J. Abd Ali, M. S. Ahmed, A. Mohamed, H. Shareef, and R. Z. Homod, “Artificial Neural Network Based Controller for Home Energy Management Considering Demand Response Events,” *ieeexplore.ieee.org*, 2018, doi: 10.1109/ICAEES.2016.7888097.

[44] A. Hussain, A. Abbas, W. S. N. Al-Rekabi, and A. N. A. Hamdan, “Prediction of Potential Electrical Energy Generation from MSW of Basrah Government,” in *Ecology and Biological Sciences*, 2017, doi: 10.15242/dirpub.er0517030.

[45] O. H. Mohammed, “Optimal Hybridisation of a Renewable System to Fulfill Residential Electrical Load: In Mosul, Iraq,” in *2018 3rd Scientific Conference of Electrical Engineering, SCEE 2018*, 2018, doi: 10.1109/SCEE.2018.8684206.

[46] K. Abd Al-zahra, K. Moosa, and B. Jasim, “A comparative Study of Forecasting the Electrical Demand in Basra city using Box-Jenkins and Modern Intelligent Techniques,” *Iraqi J. Electr. Electron. Eng.*, 2015, doi: 10.37917/ijeee.11.1.12.

[47] E. Fabbiani, A. Marziali, and G. De Nicolao, “Ensembling methods for countrywide short term forecasting of gas demand,” *researchgate.net*, 2019.

[48] D. K. Barrow and S. F. Crone, “Dynamic Model Selection and Combination in Forecasting : an Empirical Evaluation of Bagging and Boosting,” *eprints.lancs.ac.uk*, 2011.

[49] M. A. Khairalla, X. Ning, N. T. AL-Jallad, and M. O. El-Faroug, “Short-Term Forecasting for Energy Consumption through Stacking Heterogeneous Ensemble Learning Model,” *Energies*, 2018, doi: 10.3390/en11061605.

[50] H. Al-Bayaty, T. Mohammed, W. Wang, and A. Ghareeb, “City scale energy demand forecasting using machine learning based models: A comparative study,” in *ACM International Conference Proceeding Series*, 2019, doi: 10.1145/3368691.3368716.